ONE SIZE DOES NOT FIT ALL:

REGIONAL ECOLOGY, FIRM SIZE, AND INNOVATION PERFORMANCE

A Dissertation
Presented to
The Academic Faculty

by

Hsin-I Huang

In Partial Fulfillment
of the Requirements for the Degree
Doctor of *Philosophy* in the
School of Public Policy

Georgia Institute of Technology
December 2012

ONE SIZE DOES NOT FIT ALL:

REGIONAL ECOLOGY, FIRM SIZE, AND INNOVATION PERFORMANCE

Approved by:

Dr. John P. Walsh, Advisor
School of Public Policy
Georgia Institute of Technology

Dr. Janelle Knox-Hayes
School of Public Policy
Georgia Institute of Technology

Dr. Gordon A. Kingsley
School of Public Policy
Georgia Institute of Technology

Dr. Alexander Oettl
Ernest Scheller Jr. College of Business
Georgia Institute of Technology

Dr. Giacomo Negro
Goizueta Business School
Emory University

Date Approved: [July 27, 2012]

ACKNOWLEDGEMENTS

Every dissertation has a story, mine does too. It started from the January of 2007, when I entered Georgia Tech and joined in the "Walsh Lab" (A lab of John Walsh). Like many of other dissertations, my work would not exist without the help of a large number of people.

My greatest appreciation goes to my mentor, Professor John Walsh. As my advisor, John always provided me the most helpful guidance and had been the main source of motivation during my graduate program; most importantly, he kept me alive. John is a strict mentor, the haunted figure in my dreams, but he is also an inspiring mentor, a generous person who never hesitated to give me words of encouragements and support throughout. If I estimate the total amount of my two and a half years dissertation work, I am more than convinced that John was on my side leading me towards the right direction for more than 1,000 hours while constantly presenting resources and creative minds. I would also like to thank my committee members, Gordon Kingsley, Janelle Knox, Giacomo Negro, and Alex Oettl. Janelle always has constructive comments and advices in shaping my literature reviews and constantly cheering me up. I appreciate Alex and Giacomo for their friendly conversations and comments on my dissertation. Finally, to all other people and organizations that ever helped me in completing this dissertation, I would like to extend my appreciations to Jennifer Clark, Jan Youti, RIETI and the Kauffman Foundation.

My thanks also go out to Yeonji and Taehyun, my supportive labmates. I still remembered the summer of 2007 we mailed out 1,000 survey envelop a day in an assembly line. I also enjoy the friendship with Anne, who is like my older sister in the US. Thank you for being so considerate and could always brighten my day every time we met. Throughout my graduate student life, I appreciate my parents in Taipei, my aunt and uncle in Irvine, who always acted as my biggest advocates. I am also very grateful for the company of my Taiwanese folks, Ya-Lin, Cherry, Agnes and so many other TSA students, making Atlanta feel more like home.

Finally, I give my greatest thank to Victor, my best friend and fiancé, who endured my emotional swings during the dissertation writing process. In addition, I want to thank Victor proofreading and criticizing many earlier drafts. I am glad to have him alongside me at every important decision-making moment. This is not the end of the story, I will carry out everything I learned and experienced from all of you to my next destination. Once again, thank you all!

Hsini Huang

Summer 2012, Atlanta

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

This dissertation aims to answer the main question of "How does regional ecology (few or many small innovative firms in a region) enhance or limit innovation?" Put differently, how vital is the mix of small and large firms for regional innovation performance? From the policy perspective, the results of this study shed some light for policy maker to assess the "knowledge searching" strategies of firms when choosing locations. The research design combines a unique survey of patent inventors in the United States and archival data. Georgia Tech inventor survey data contains commercialization measures for patented inventions and information on firm characteristics. Using this archival data, data has been collected on regional innovation measures, regional-level attributes and project-level measures.

The results indicate that the agglomeration of specialized firms is positively associated with regional innovation activities, as the Marshall-Arrow-Romer model proposed. In addition to traditional regional measures, small firm dominated ecology is a strong factor explaining regional commercialization activities, even though the role is not very significant when explaining the regional patenting activities. It is suggested that the organizational ecological perspective is complementary to understand information flow mechanisms in innovative regions. One mechanism of SME dominated ecologies is partially through the increase of skilled labor mobility. Furthermore, when the regional ecology moves towards being dominated by small firms, large firms benefit more from the presence of many innovative small firms than SMEs. By contrast, the concentration of innovative small firms does not add much value for SMEs. I suggest the focus of policies should be on understanding the heterogeneous ability of accessing localized knowledge resources between large and small firms. Deriving from the findings, policy implications and future research are discussed.

CHAPTER 1

INTRODUCTION

1.1 Introduction

1.1.1 Research background and policy implications

Innovation is a key feature of competitive advantage for firms to continue growing in the knowledge-based economy. Firms increasingly rely on integrating external knowledge with their existing capabilities in order to achieve successful R&D and innovation. Similarly, innovations take place in the context of an environment and are the result of the interactions between players in the same innovative system. The idea of regional economies has been picked up by the federal government as well. In September 2011, the Obama government started a new program, the Job and Innovation Accelerator Challenge (JIAC), focusing on the development of regions' innovative ecosystems. The purpose of this program is to promote regional innovation clusters and increase jobs.

However, under the current economic policies, the concept usually concerns establishing a general environment that affects all firms (e.g., the federal tax benefit for a particular industry), or to allocate resources to certain individual firms (e.g., the JIAC program and the Small Business Innovation Research (SBIR) grants). The pitfall of these "one size fits all" policies is the disregard of the discrepancy between the national/state's average standards and the regional goals. At the regional level, the development of clusters tends to largely emphasize the scale of the economy, but not the intra-regional structure. This research aims to fill this gap. Therefore, before discussing the implementation of a regional level economic policy, it is proposed that we should first understand the environment and organization relationship in a region and its effect on innovation performance.

The major theme of this research is to realize the role of regional ecology on innovation performance. Regional ecology is defined as the distribution of large and small firms in a region. A recent study by Clark, Huang, and Walsh (2010) illustrates the significant variation in rates for small firm patents across metropolitan areas in the United States. One question that is important to ponder is whether types of regional ecology explain the variation in regional innovation performance. In particular, does the concentration of many innovative small firms provide a more sustainable innovative region? This current research takes an organizational-ecology perspective to explain the impact of firm colocation on innovation performance. It aims to understand how different organizational ecologies shape regional structures to enable interactions among firms in the same region. This study also explains how the regional ecology influences the circulation of skilled labor and local knowledge, and further affects inventing activities and the pursuit of commercialization.

A second theme concerns discussing how vital small firms are to the regional economy. In both political and policy debates, the argument that small firms are the backbone of our national economy for both job creation and innovation is commonly made (Obama, 2009, c.f. Clark et al., 2010). Some argue that young and startup firms are particularly valuable (Haltiwanger, Jarmin, and Miranda, 2010; Delgado, Porter, and Stern, 2010). According to a Small Business Administration (SBA) report in 2008, small and medium sized firms accounted for 69% of the (non-farm) net new jobs from 1993 to 2008 (SBA, 2010). United States Patent and Trademark Office (USPTO) statistics also indicate that small firms are sharing more than one-third of the issued patents in 2000. However, this number gradually decreased in the 2000 (35%) - 2009 (28%) period. All these numbers show that small firms substantially contribute to national economies. The question is whether innovative small firms can receive resources from the locality they need to be competitive in innovation.

This research provides insights into how firms benefit from their geographic locations. From a policy perspective, the results of this study could illustrate a better framework for policy makers to assess the "knowledge searching" strategies of firms when choosing locations. Prior studies suggest that firms are required to actively search for local knowledge in order to achieve innovation (Jaffe, Trajtenberg, and Henderson, 1993; Zaheer and Hernandez, 2011). Therefore, firms tend to choose locations for gaining potential knowledge spillover sources (Alcacer and Chung, 2011) to maximize the inflow of knowledge. In addition, it is important to tie the spillover argument with firm heterogeneity in innovation creation, knowledge sharing and knowledge appropriability. The main argument in this study is that large and small firms are facing different complimentary constraints when conducting R&D and innovation activities. In the population ecology theory, the resource-partitioning theory suggests that members in a population are likely to compete over finite resources. The intensity of competition between organizations in a population is a function of their similarity for resource requirements. In other words, the more similar the resources are, the greater the potential for competition (e.g., McPherson, 1983). Specialized firms are likely to partition the resources in a concentrated market. On the other hand, the niche-width theory suggests that generalists could perform better in a fast changing environment. These two theories indicate the constraints that firms face both spatially and ecologically.

If differently sized firms carry heterogeneous capacities, firms may not benefit equally from the geography in which they are located. Concerning regional development, one universal economic policy might not apply to all regions, or to all firms in a region. Current policy practices tend to focus more on the aggregated outcomes in a region, such as aggregated innovative activities and overall employment growth. However, firms often merely emphasize individual benefits but not collective interests. This study argues that tensions may exist between the regional policies and individual firms towards an effective innovative region and innovative opportunities of firms. Policymakers for regional development should consider whether location and its organizational composition benefit differently sized firms.

1.1.2 Firm colocation, regional ecology, and innovation

The study of regional economies and its role in production and innovation is not new. The major literature focuses on the theory of agglomeration economies, emphasizing that the spatial concentration of firms generates external effects to firms in the similar industry (Marshall, 1920; Porter, 1990; Glaeser et al., 1992; Feldman & Audretsch, 1999; Feldman and Kogler, 2010). Agglomeration theory implies positive external effects for the collocation of firms and is associated with the capital returns to firms located in a region (Bresnahan and Gambardella, 2004). These economic external effects are embedded in the space that directly and indirectly facilitates the growth of the region (Cooke and Morgan, 1994). For example, the advantages of firm concentration are a reduction in transportation costs, sharing of infrastructure, accessibility to a large pool of skilled labor and the sharing of ideas. In addition, the concentration of firms can increase the influx of specialized suppliers, as proposed by Marshall (c.f. Stuart and Sorenson, 2003). In other words, previous industrial cluster research considered the geographic proximity as a stand-alone determinant to explain manufacturing and knowledge production. The collocation of firms increases the likelihood of sharing information and ideas with colleagues in neighboring firms, and further increases the chance of discovering new technologies and innovations.

The contribution of this current research is to expand agglomeration theory by proposing that we should not only consider the effect of firm concentration, but also the types of concentration (i.e., the regional ecology) in relation to innovation performance. This study proposes that the conceptualization of regional ecology helps us better understand the innovation process for both firms and regions. Deriving from population ecology theory and industrial district theory, the regional ecology is defined as the mix of large and small firms in a region. The regional ecology represents an environmental context, which is different from the traditional concentration indices in previous regional studies. The concept of regional ecology will be constructed by measuring the distribution of innovative activities by firm size for each region. From an ecological perspective, the regional structure of size concentration can explain part of the organizational constraints in the conduct of innovation.

In summary, this research aims to answer the main question of "How does regional ecology (few or many small innovative firms in a region, as an inverse measure of few or many large innovative firms in a region) improve or limit innovation?" Put differently, how vital is the mix of small firms and large firms for regional innovative performance? This dissertation addresses the question of whether the innovation performance of small and large firms in the United States is influenced by the regional-level ecology when using firm-level characteristics as the control.

Following Schumpeterian tradition, "innovation" defined in this study has two parts, invention and commercialization. Invention refers to novel and useful technologies. Commercialization refers to the commercial use of new technology (radical invention) or a new combination of existing technologies (incremental invention) (Afuah, 2003; Schumpeter, 1942; Jung, 2009). The process of invention and commercialization are both important to economic growth, while the effect of regional ecology may play differently at the R&D stage.

1.1.3 Research questions and analysis

This research asks three research questions. First, how does agglomeration (regional resources) affect innovation? Secondly, does regional ecology enhance or reduce innovation? Finally, how do the effects of regional contexts, i.e., the regional ecology effects, differ between large and small firms? Put differently, do firms benefit from the location and who benefits more?

To answer the above research questions, two sets of empirical analyses will be used. The first one analyzes regional factors that determine innovation performance at the regional and firm level. The analyses will be conducted in the following order. First, we will look at the impact of regional resources (i.e., university resources, labor mobility, and specification) on regional development by examining whether agglomeration increases innovation performance. We will then examine whether regional ecology, the distribution of types of organizations in a region, influences the innovation performance of regions and firms. These analyses examine the innovation outcomes of R&D projects as a function of regional resources and regional ecologies, controlling both firm and project characteristics.

The second set of analyses examines how the external effects of regional ecology differ by firm size. In other words, do regional ecologies play different roles for large and small and medium-sized firms? In addition, if regional ecology represents a social structure that facilitates knowledge sources, then to what extent are the effects of ecological contexts (regional ecology) mediated by different knowledge flows mechanisms (the regional knowledge sources)?

The research design combines a unique survey of patent inventors, the RIETI (The Research Institute of Economy, Trade and Industry)/GT Inventor Survey (N = 1,919) in the United States and several pieces of archival data (e.g., bibliometrics patent documents (PATSTAT), and census statistics). The GT inventor survey data is the major dataset, containing commercialization measures for the patented inventions and information on firm characteristics. Using archival data, I collect data on innovation measures, firm-level characteristics, the regional ecology measure, regional mobility rates of skilled workers, and university R&D expenditures. For the analyses, both OLS (Ordinary Least Square) and HLM (Hierarchical Linear Model) regressions will be used to account for regional (MSA), firm, and project level effects.

1.1.4 Dissertation structure

This dissertation consists of five chapters. The remainder of Chapter One presents the theoretical framework by reviewing the existing research and literature on organization theory, economics of innovation, and economic geography to identify theoretical gaps between these fields. In particular, Chapter One reviews the literature that examines the interplay between firm size and regional characteristics and its influence on innovation performance. It constructs the concept of "regional ecology" as the key theoretical contribution. It also derives testable hypotheses based on the synopsis of the existing empirical and theoretical findings.

Chapter 2 continues by presenting the methodology, study design, data source, and analytical strategies used in this study. Chapter 2 also includes a section describing the limitations of the data. Subsequently, Chapter 3 analyzes the impacts of agglomeration effect and regional ecology on innovation performance at both the regional and patent levels, thereby. It answers the first research question ("*do regional resources (e.g., specification of industry, university knowledge, and regional labor mobility) improve or limit regional innovation performance?*") and the second research question ("*does regional ecology, i.e., few or many small innovative firms, improve or limit regional and firm innovation performance?*"). Chapter 4 addresses the third research question, "*How does the effect of regional ecology differ by firm size?*" This chapter tests whether small firms are able to benefit more from a small-firm dominated region (referred to as a Marshallian thesis) or a large-firm dominated region (referred to as an Anchor-tenant thesis). In addition, Chapter 4 also analyzes whether firm size moderates the effect of regional ecology on the innovation performance of regions and firms.

In concluding the research, Chapter 5 combines the findings of Chapter 3 and 4 to provide conclusions on the relationship between regional ecology, firm size, and innovation performance. Chapter 5 also discusses and compares findings of this research with prior literature, particularly the literature reviewed in Chapter 1, as well as providing policy and managerial implications for policymakers of regions and firms. Based on the findings, the concluding chapter also discusses future research areas.

1.2 Literature review

The following sections summarize and elaborate on the existing economics of innovation, and economic geography literature. Based on current literature debates, a series of research questions and testable hypotheses will be constructed. The structure of this chapter focuses on two theoretical themes, 1) Agglomeration, knowledge spillover, and innovation, and 2) regional ecology and innovation. It begins by discussing innovation to introduce the background of this study. Then, the section discusses firm size and firm level determinants used when conducting R&D and innovation. Secondly, before discussing the regional ecology concept, we will review seminal agglomeration theory to bring out the concept of firm collocation and its impact on knowledge spillover and regional innovation performance. Finally, by summarizing previous literature on industrial districts, we will introduce the need for an ecological perspective to understand the region-organization relationship and its impact on innovation performance, naming the regional ecology. It will also review empirical papers that discuss the inside structure of regions, as well as timely discussions about the interplay of regional ecology, firm size, and innovative performance.

1.2.1 Why study innovation?

Innovation as technological change has a positive destructive effect on the economy. However, innovation is a complex and institutionalized process. This study defines innovation as both invention and commercialization. Invention is novel and useful technologies. Commercialization means the commercial use of new technology (radical invention) or a new combination of existing technologies (incremental invention) (Afuah, 2003; Schumpeter, 1942; Nelson and Winter, 1982; Jung, 2009). The theoretical root goes back to Joseph Schumpeter (1942), who emphasized the relationship between market structure (associated with firm size) and innovation. Schumpeter defined "innovation" as the actual introduction of the novel inventions, such as new processes, new products, new materials, or new services. Schumpeter not only argued the possibility of entrepreneurship in innovation, but also the role of monopolization in innovation. Concerning entrepreneurial activity, the encouragement of entrepreneurships provides the possibility of the destruction of social status, and the reordering of the economic system. He proposed that entrepreneurial activities are the foundation of the competitive market for innovative knowledge and technology. For the latter one, monopolization refers to a large firm possessing the advantages of better resources and financial standings. In sum, Schumpeter's theory introduced the potential of innovation creation as the new page in the capital system. He implied that differently sized organizations are under different conditions of competition when doing innovative business.

Kenneth Arrow's (1962) innovation model suggested that the value of the invention is its information. Patent system could be an appropriation path through which to protect the incentive of inventors because information is by nature a non-rival good. He implied that knowledge developed for any inventor could easily spill over to other firms (c.f. Feldman and Audretsch, 1999). For example, as Arrow mentioned, "Mobility of personnel among firms provides a way of spreading information…" Similarly, Nelson and Winter (1982) suggest that innovation is an outcome of organizational learning. They provide elaborated analogies explaining why a successful technological change requires a search for knowledge and the selection of an appropriate environment.

According to Nelson (2001), the selection mechanism operates within a firm's boundaries and refers to the behavioral and technological options that are selected and retained by firms from their available resources. Nelson and Winter's theory is somewhat drawn from a "biological conception" that organizational learning and innovation is a socially structured process.

To measure innovation activities, previous studies emphasized firm level R&D outputs (e.g., the Carnegie Mellon Survey of Industrial R&D of manufacturing sectors in 1994; the Community Innovation Survey, 1995; 2000), particularly on the counts of innovations (Acs et al., 2002). Patents were also used as a proxy for innovation activities in related studies. Patents represent an intermediate measure that is better than the R&D expenditure measure because budgeted resources are not necessary equal to performance (Griliches, 1991). However, patent documents are longitudinal data and openly accessed to the public, hence more and more scholars use bibliometrics data to construct innovative performance measures. Acs and his colleagues (2002) find that patent counts could be a reliable measure of innovative activities because regression outputs were similar with results for predicting innovation counts. Other scholars use patent inventor survey data to investigate the economic and technological value of patents, as well as the process of patent commercialization (Macdonald, 1986; Mattes et al., 2006; Gambardella et al., 2008; Nagaoka and Walsh, 2011). Compared to bibliometrics patent data, the main advantage of surveying patent inventors is to obtain detailed information about the innovation process. Survey data allows us to ask the inventors about the R&D process and the use of the invented technology at the time they were involved in the patented project. Previous studies suggest that we should not only explore the overlapping concepts of patent, invention, and innovation but also push the research question forward by asking what transforms an invention into a commercialized innovation. It is likely that an invention needs not to fulfill customers' needs and requires less concern for the exploitation of the concept in the marketplace. Therefore, an invention can be measured by its patenting propensity. In contrast, a commercial innovation requires matching with certain market demand. Hence, the drivers of commercialized innovation can be different from the drivers of inventions.

In summary, bibliometrics patent data can be useful in identifying inventions with potential appropriate value because filing patent applications requires a lot of time and money. Surveying patent inventors has the advantage of being able to trace the innovation process from its R&D phase through to the commercial use phase, and to control certain project level characteristics, such as the scale of R&D inputs.

Firms and Innovation

The following section will first discuss drivers of innovation for individual firms. In the knowledge-based economy, innovation is a key to economic growth (Nelson and Winter, 1982) and competitive advantages. In the past fifty years, to understand the process of innovation, organization theorists have been dedicated to studying the drivers of innovation activities by firms. One key determinant is firm size. First, Schumpeter claims that the innovation performance grows disproportionately as the size of firm increases. This assumption suggests that larger firms are more likely to conduct and invest in R&D activities than small firms, therefore while firm size is controlled, either R&D inputs or outputs are positively associated with innovative quantities (Cohen, Levin, and Mowery, 1987; Pavitt, 1991). However, the counter-argument is that small firms' innovative performance per unit of R&D inputs is greater than that of large firms (Audretsch and Acs, 1991; Pavitt, 1991). This implies that small firms tend to choose R&D projects that are more likely to be applied and commercialized. Another reason is that small firms outperform large companies because the diminishing productivity of R&D is more obvious for large firms when the productivity of every additional investment in R&D dollars is decreasing (Cohen and Klepper, 1996). However, if we assume the diminishing productivity of R&D is equal for large and small firms (they are both efficient) then large firm should benefit from not betting all their money in a few projects since the overall productivity should be higher than that of the small firms.

The other mostly discussed determinant is firm capability. Teece (1986) argues that the possession of complementary capabilities (e.g., manufacturing facilities, services, and complementary technologies) is needed for commercialization (Teece, 1986). Awareness about unobserved firm heterogeneity has been raised by a new influx of studies, attempting to measure the R&D capability, such as absorptive capability (Cohen and Levinthal, 1990), the ability to acquire and use knowledge, or the capacity of managerial resources (Kremp and Mairesse, 2004).

Much research has explores the relationship between the size of firm and the capacity of production and R&D assets. Large firms are more likely to possess greater complementary capabilities than small firms. For example, large firms benefit disproportionally more from advanced knowledge from university research compared to small firms because of a greater amount of PhD degree graduates being hired (Cohen, Nelson, and Walsh, 2002). In summary, prior work suggests that firm characteristics are a key predictor of invention and innovation. However, these firm-level effects need to be put into context. Starting in the next section, this study takes the multidisciplinary approach and draws from literature on the economics of geography to investigate the relationship between environmental contexts and innovation performance.

1.2.2 Space and innovation

To understand the relationship between space and innovation, this section reviews the literature on agglomeration theory, learning region theory, industrial districts theory.

1.2.2.1 Early theories on agglomeration

The seminal work of Marshall's (1920) agglomeration theory emphasizes that the advantages of geographic proximity not only reduce transportation costs, but also help learn new skills from neighbors and assure a constant supply of labors. Marshall clearly identifies three important resources gained from the concentration of manufacturing firms, including transportation facilities, skilled workers, and ideas. Marshall (1920) suggests that the concentration of specialized firms enjoys similar economies of scale a large firm (Marshall, 1920, IV.X.21). A group of specialized firms can therefore expand/grow in a particular place because of the use of external economies.

Later scholar developed the Marshal-Arrow-Romer (MAR) externality model (Glaeser et al., 1992) that proposes that the concentration of specialized industries positively associate with inter-firm knowledge spillovers in a particular region. The MAR model claims a specialized region could grow faster for two reasons. First, local concentration increases within-industry knowledge flows. Secondly, the concentration of firms in the same industry increases local competition among firms, resulting in more incentives to innovate.

In contrast, Jacob (1969) proposes that the concentration of diversified industries is better for regional growth because of the cross-fertilization of ideas across different industries, resulting in unexpected new technologies or services. For example, many banking services and products were not invented by the financial sector, but by ancillary industries or users. Empirically, in the U.S., cities with specialized industries were decreasing employment growth (Gleaser et al., 1992). In Gleaser's study, specialization is a measure of the concentration of a particular industry in a city. Additionally, they find that the numbers of firms per worker in those city-industries with high growth rates are larger than the national average. For example, firms in the electric machinery industry collocated in San Jose, California are smaller than the national average size of firms of that industry. Glaeser's findings are similar to Jacob's argument concerning important knowledge potentially coming from outside the core industry, rather than within the industry. Moreover, they suggest a positive correlation between the local competition and the regional growth. In summary, agglomeration theory emphasizes either the industrial homogeneity or heterogeneity in contingent with the concentration of firms.

In the 1990s, the increase in global trade and the development of new information technologies reshaped the global economic landscape. However, globalization is not geography-free. The concentration of production and new financial services are clustered in a few global cities (Sassens, 2001; Dicken, 2003. For example, the financial service market became more clustered in the global cities (Sassens, 2001; Clark, 2002), particularly in London, New York, and Tokyo. This phenomenon suggests that face-to-face communication and social ties within physical distance are important to the knowledge-based industries.

Combining physical proximity and knowledge economy perspectives, Florida (1995) suggested that a “learning region” could be sustained in a new era of capitalism. Different from traditional manufacturing regions, learning regions constantly supply infrastructures that facilitate manufacturing, human resources, communications, and industrial governance systems that are required for knowledge-intensive economies. Florida emphasizes that knowledge is the essential component for innovation in the learning region. Built upon the assumption that geographical proximity facilitates knowledge spillovers, the “learning region” argument is consistent with Romer’s (1986) claim that “knowledge spillovers” are the engine of economic growth.

Some argued that the link between geography and regional technological growth is more than physical convenience. The institutional structure within a region explains why some regions can bring in localized advantages, while others cannot (Saxenian, 1996). Saxenian compares Silicon Valley in California with Route 128 in Boston to illustrate the formation of a successful technological region. The electronics industry in the Route 128 region around Boston began to grow because of a vast influx of government defense funding from the 1960s. Although the major proportion of money went to large companies, such as DuPont, Kodak, and Xerox, some new computer companies, such as DEC and Lotus Development, were funded to provide complementary services. However, Route 128 did not maintain its advantages too long. Since the 1980s, Silicon Valley overturned the leading position of Route 128 in electronic industry. Saxenian argues that Silicon Valley had a very different social and cultural structure than Route 128. The dense regional network, flexible institutional culture, and positive loop of mobility within the Silicon Valley led to its prominent success. However, some scholars argue that the growth of Silicon Valley was actually led by legal differences between California and other states, particularly the enforcement of non-compete clauses between the two regions (Gilson, 1999). Empirical studies found that the enforcement strength of the non-compete clause decreased turnovers and spin-outs, and ultimately reduced new innovative entries (Fallick et al., 2011; Garmaise, 2009; Marx et al., 2009; Singh and Marx, 2011).

To summarize, the agglomeration economies explain the external resources could be accessed by firms collocated in the same region. The Marshall tradition emphasizes the advantages of geographic proximity and the influx of many specialized firms. The development of theories, such as learning regions and territorial innovation systems, has gradually shifted from a discussion of collocation of producers (often connected through value-chains) to the collocation of innovators (Simmie 2005). Florida and many learning region scholars address the locus of knowledge to innovation and regional sustainability. Saxenian brought up an interesting discussion regarding the role of institutional structure and regional culture in shaping the inter-firm interactions and related regional resources.

As Feldman (1994) summarized that the collocation of firms is important to innovation in providing the following regional resources. First, collocation facilitates the concentration of information and knowledge resources (von Hippel, 1988). Secondly, collocation increases the chance to acquire university research for the local (Dosi, 1988). Thirdly, collocation reduces the uncertainty when undertaking the innovation (Dosi, 1988). Finally, collocation carries pools of technologies, skilled labor, and cumulative knowledge (Saxenian, 1996; Powell, 1990). In other words, the concentration argument considers geographic proximity as the key venue for knowledge spillovers within the industry or across industries. However, while previous literature is heavily industry-oriented, we argue that to understand regional development, we should initially take organizational heterogeneity into account. Therefore, the organizational ecology perspective is complementary when understanding the complex mechanism of knowledge flows in a region and its impact on regional innovation performance.

1.2.3 Ecological perspectives of firm collocation

In this research, the concept "regional ecology" is borrowed from Hannan and Freeman's population ecological perspective (1977) that addresses organization-environment relations. Organizations face constraints on the information and resources they receive, and the information and resources that are available for sustenance in the environment. According to geography literature, the theory of industrial district also highlights the structure within an industrial district and its impact on regional development (Markusen, 1996; Gordon and McCann, 2000). By the end of this section, I will conclude the summary of these theories by introducing the concept of regional ecology as a structural variable, the core theme of this research.

Population ecologies

Some organization theorists view environment as the source of innovation adoption because organizations should match their capability to the environment they face. For instance, changes in work are often the results of environmental pressures on organizations that reflect the technological shift of an industry (Walsh, 1993). The population ecology theory first developed by Hannan and Freeman (1977) addressed that the resources a firm can access are constrained by the population of organizations where they are located. They study the performance of organizational ecology by measuring the birth and death of firms in an organizational population. Their theory is based on the following observations. First, aggregates of organizations exhibit different levels of diversity. Secondly, organizations have difficulty adjusting to changes fast enough to meet the demands of uncertain and environmental variations. Finally, organization populations evolve (enter and leave) continually. Therefore, to study organizational growth, the “population” should be the unit of analysis clarifying the association between organizations and the environment.

Hannan and Carroll presented two important theories in the 1980s. One is Niche-width theory that refers to the variation of organizational strategies can be utilized in an environment with defined scope of resources. The major question is how environmental dynamics affect the niche width of a certain population group (Popielarz and Neal, 2007). Types of organizations—generalist and specialist—need be considered when studying the ecological impact. Carroll and Hannan (1977, 2003) imply that organizations seek regional resources to sustain themselves. To summarize, it argues that specialists are betting all their resources (technologies, in this study) on specific outcomes, while generalist organizations hedge their inventions. In other words, generalists take less risk than specialists do when the environment changes because they tend to distribute their investment in many different areas. On the other hand, specialists are in the advantageous position to cover the narrower niche in a stable environment.

The other one is the resource-partitioning model that explains how the local resources realized by individual firms in different environments. The focus of this theory is on answering the partitioning of two non-competing populations in the market (Carroll, 1985; Popielarz and Neal, 2007)

Empirically, they found that small specialists are able to survive in the presence of large generalists by finding a special segment of customers, like in the newspaper (Carroll, 1985), and the beer brewing industries (Carroll and Swaminathan, 2000). The resource-partitioning model has certain theoretical assumptions. The assumptions are, 1) organizations have limited ability to adapt to environmental changes, 2) organizational choices are constrained by bounded rationality, 3) the market contains finite resources drawn by organizations, 4) no price competition among firms exists in the environment, and 5) consumers in the market are heterogeneous. In the newspaper industry, Carroll found that specialists would exploit more resources from the environment than generalists would since specialists could draw more resources from a concentrated market without competing with the generalists directly. On the other hand, generalists face higher mortality rates than specialists in a concentrated market because the generalists are not able to occupy the peripheral niches of the small specialists. In summary, the Resource-partitioning perspective and the Niche-width theory describe the constraints of organizations collocated in a region. Organizations have to seek for their niches and organizational strategies in response to the resource-space in which they are located.

This research applied the concept of population ecology theories to firms in the innovation business. The technology markets in high-tech industries (either patent-based or non-patent-based) are likely to follow those important conditions Carroll mentioned. First, high-tech firms tend to be clustered because they rely largely on local human capital and local finance, particular high-tech entrepreneurs. Secondly, the resources are heavily concentrated in the center of the market, such as R&D-rich large corporations.

By referencing the concept of organizational ecology developed by Hannan and Carroll, this current research adopts the idea of finite resources in an environment in which specialists and generalists are having distinct niches to survive when under the concentration of large (generalists) firms. The research setting does not try to corroborate Hannan and Carroll's theory, but the idea is to investigate whether specialists are likely to find their niches when there is the concentration of many small firms, suggesting that innovation space is not dominated by one or a few large firms. This study explores the organizational ecology at the regional level and will measure the distribution of firm size in a region and investigate its impact on the population performance.

Similar to the ecology population perspectives, Feldman and Kogler also state that "…while firms are one venue to organize economic activity, the resources required to generate innovation are typically not confined to single firm, and geography is another means to organize the factors of production" (Feldman and Kogler, 2010**,** p404). In this study, one major difference is the use of invention counts and commercialization rates as measures of performance of organizational populations.

Industrial Districts

For economic geographers, the industrial district theory was born from observing a unique regional structure based on successful stories in the north-central and north-east region of Italy (Harrison, 1992; Piore and Sabel, 1986). They mostly focused mostly on clusters of small family firms led to successful manufacturing production. Markusen's (1996) contribution to this thread of theory was to develop a typology identifying four distinctive types of industrial districts. Markusen illustrates the diversity of spatial clusters and provides insights into structures of industrial districts (e.g., connections between differently sized firms). Based on qualitative surveys of regions with superior growth rates since the 1970 in the United States, Japan, South Korea, and Brazil, Markusen defined sticky (i.e., successful) industrial districts (ID) as the following. (a) Marshallian ID: A region is comprised of a large percentage of specialized small firms and significant levels of local networks and cooperation among small firms. (b) Hub-n-Spoke ID: A region dominated by one or a few large firms that are heavily engaged in the local economy, with the presence of a dominant large firm also meaning a domination of one or a few industries. (c)Satellite Platform ID: A region dominated by branches of large corporations with employees committed to firms but not to the district. (d) State-Anchored ID: A region dominated by government institutions, such as government laboratories, military bases, or universities.

To describe the processes of the different types of regional structure in more depth, Gordon and McCann (2000) studied industrial districts in London and distinguished three typical industrial district models, including the pure agglomeration model, the industrial-complex model, and the social network model. The pure agglomeration model is similar to the Marshallian district of Markusen's typology in which the concentration of many small and medium sized firms brings many advantages and inter-firm learning. The externality of agglomeration does not require firms co-located in the same region to have intensive interactions. In the industrial-complex model, those key players are often large in scale and seeking for profit monopoly. The third model is the social network model that suggests a more integrated community among partners in the industrial district, for example, the Silicon Valley story by Saxenian (1994), and the Hollywood story by Storper (1997).

In sum, Markusen reminded us about looking carefully at the internal structure of a region by observing the connections between different types of firms. However, Markusen's research had the following limitations. First, to summarize the typology theory, her research design heavily depends on a few successful cases. We do not know if this typology described less successful regions or not. Secondly, the definition of a prosperous region was documented mostly on the traditional measures of manufacturing productions (e.g., employment growth and manufacturing change). There was little emphasis on the connection between the innovative activity and regional structure.

This section bridges two sets of literature, the population ecology theory from the organization theory, and the industrial district theory from the economics of geography. The literature review shows that the effects of environmental contexts on organizational structure and regional structure are not trivial. Hence, my study proposes investigating regional economy from an ecological perspective and to understand how regional ecology plays as an environmental driver of innovation. What is still interesting in existing research is what types of regional structure contribute to the innovative performance of firms and regions.

In addition, my study uses the framework of Markusen's typology of industrial districts but with two important modifications. First, this study defines the "regional ecology" as mixed firm size in a region to represent the firm size composition across regions. The regional ecology is a relative concept because it can be structured by a large percentage of local small firms, or a few major large corporations that dominate the majority of innovative productivity. The second modification is the use of innovation data rather than employment data as the key regional indicator.

The following sections will now review literature on studies that discuss internal regional structure, which is similar to the regional ecology concept proposed above. In addition ,it will also review previous studies that discuss the influence of regional structure on innovation, as well as the hypotheses of this study.

1.2.4 Regional ecology and innovation

In this study, regional ecology refers to the structure of the size concentration of firms in a region, particularly the distribution of types of firms in each region. The following section summarizes related literature discussing the relationship between the regional ecology and innovation performance. One type of ecology is small firm dominated ecology (the Marshallian district, Markusen, 1996). The second type is large-firms dominated ecology (the Anchor-tenant region, Agrawal and Cockburn, 2003; Markusen, 1996). Some recent studies found mixing of a few large firms and many small firms, providing a hybrid environment that increases the regional innovation performance (Agrawal, Cockburn, Galasso, and Oettl, 2011). The role of a small firm dominated ecology is particularly interesting and consistent with the current policy focusing on creating high-tech regions with a cluster of many innovative small firms. Some studies also view the presence of many small firm innovators as a proxy for "embedded institutional capacities," which enhances regional long-term growth and resilience (Clark, Huang, and Walsh, 2010).

Small-firm dominated ecology and regional innovative activity

In the late 1990s, neo-Marshallian theorists revisited Marshall's agglomeration theory and emphasized the role of co-operation networks among small firms as the driver of successful regions. For example, high-tech regions in the United States and sustainable craft-based industrial districts in the Third Italy. This stream of theory emphasized on the advantages of flexible specialization and lean production of small firms in the post-Fordist era (Piore and Sabel, 1986). A recent empirical study shows that Marshallian-like innovation districts in the U.S. have higher GDP per capita than other types of districts (Clark et al., 2010). The advantages of being in a small firm cluster are several. First, the concentration of small firms in the same industry reduces transaction costs and increases untraded interdependencies, such as the film making industry in Hollywood (Storper, 1997). Secondly, the agglomeration of many small firms can enhance the complementarity advantage of firms collocated in a region. The concept of complementarity means that each institute provides a special kind of service in the region. For example, institutions and intermediaries of London's financial industry not only compete, but also complement each other based on functionality (Gordon Clark, 2002). Similarly, the existence of small firms provides complimentary services in diversified areas, which are less likely to be provided by large firms. Thirdly, the collocation of small firms in a district/region creates collective advantages, flexibility and specialization in particular (Pyke and Sengenberger, 1992), which is positive for local competition. Collective efficiency allows specialized small and medium size firms to catch up with the technologies of large firms. Related to this argument, the concentration of specialized firms also implies an increase in the diversity of technological knowledge domains even among firms in the same industry because an innovation often comes from the combination of existing ideas and technologies (Fleming, 2001). Fourthly, trust is the important adhesive byproduct in the Marshallian district and essential for establishing long-term relationships and a dense social network. As Owen-Smith and Powell (2004) argued, a dense social network among regionally agglomerated firms reduces the risks of opportunism and creates signaling effects among members in the same network. As a result, information is transmitted easily throughout the network.

Theoretically, a small firm dominated ecology should benefit all firms in the locality, whether large and small firms and their capacity for innovation. In a small firm dominated ecology, small firms are less likely to be a dependent of large firms. Second, large and small firms are presumably having equal ability to enjoy the aggregated knowledge spillovers if they belong to the same local network. A cluster of many small firms increases competition and ideas of new technologies. Empirical data from Small Business Administration, Acs, Anselin and Varga (2002) shows that the concentration of large firms in a metropolitan statistical area is likely to lower the regional innovative activity.

Based on the advantages of a small firm dominated ecology mentioned above, this study predicts that a SME-dominated region could outperform other types of region because it facilitates the formation of a dense network among collocated firms, with a dense network being the key to a sustainable productive and innovative region. Hence, the theory suggests:

Hypothesis 1a: As the proportion of small firm patents in a region increases, regional innovating activities (patents per capita) increase.

Hypothesis 1b: As the proportion of small firm patents in a region increases, regional commercialization rates increase.

Mechanisms of knowledge flows in a region

The previous section summarizes that the collocation of firms creates an agglomeration economy and thus facilitates knowledge spillovers in a region. This section continues that venue by further discussing the "how" question. Studies on the relationship between knowledge spillovers and regional growth are ample, with many scholars conducting research that models the role of knowledge spillovers within a geographic boundary on growth, such as Griliches (1979) and Glaeser et al. (1992), particularly in relation to employment growth and production growth.

Geography scholar like Boschma (2005) suggest that innovation has a relation with place because "diffusion" as one important outcome of innovation. To spread new technologies, ideas, and concepts, physical proximity becomes an essential issue (Glaeser et al., 1992), particularly in the early stage of technology development. Close proximity provides advantages to transmit tacit knowledge, new ideas, and interpretation of codified knowledge effectively within a geographical boundary (Audretsch and Feldman, 1999, 1996). Many empirical studies have pointed out that R&D spillovers are the reason why innovation activities were clustered spatially, especially in knowledge-based industries, like the specialized financial services in Feldman's (1994) study. Similarly, Jaffe et al.'s (1993) experiment shows that patent citation analysis can be used to trace the knowledge flows of firms. Their results confirmed a higher intensity of the citation activity and spillovers of R&D labs if they are geographically or technologically more concentrated. Jaffe, Hall, and Trajtenberg's patent citation method has become the standard procedure to examine the knowledge spillovers among firms. Later empirical studies adopted their methodology to replicate (Hicks et al., 2001), or to criticize (Thompson and Kean, 2005) the implication that knowledge spillovers are geographically bounded.

Organization theorists also view geography as a vehicle of knowledge spillovers. For evolutionary economists, organizational learning is path-dependent and heavily based on prior knowledge (Winter and Nelson, 1980). In addition, Storper and Venables (2004) suggested that innovation is a collective process through communication among inventors, entrepreneurs and other local actors. Hence, the following sections will review literature on the sharing of tacit and codified knowledge, its relationship with both firm and regional innovation performance.

1.2.5 Regional knowledge resources and regional ecology

The idea of knowledge spillovers concerns the dissemination of knowledge, with types of knowledge mattering. Prior studies categorize knowledge into tacit and codified knowledge. Nelson and Winter (1982, p. 79) defined tacit knowledge as "a part of skills that is imperfectly assessable to conscious thought." In contrast, codified knowledge means a set of skills formulated with written instructions, such as computer programs or chemical formula (Polanyi, 1967, c.f. Nelson and Winter, 1982). Nonaka and Takeuchi (1995) make a clear distinction between tacit and codified knowledge and how these two kinds of knowledge have interwoven for technology development in Japanese cases. They observed that firms could transform tacit knowledge into explicit knowledge by being close to the knowledge source. One famous example is the development of an automatic home bakery machine. The software engineer in the Japanese company had to learn the tacit knowledge about how to knead bread dough by hands from a bread master before designing the machine (Nonaka and Takeuchi, 1995). Their argument relates effective knowledge transmission with physical proximity and the outcome of the technological inventions.

For firms, compared with the traditional vertical integrated model, the open innovation model is a more efficient way to speed up the innovation process (Chesbrough, 2003). Except for the internal knowledge reservoir, firms can capture new ideas and external knowledge through different approaches. For example, the use of public literature (Cohen, Nelson, and Walsh 2002), forming strategic alliances for joint investment (Rothaermel and Deeds, 2004), joint patenting (Hagedoorn, 2003) and collaborating with universities and government labs (Powell et al., 1996) to better understand basic knowledge insights and cutting-edge science discoveries. As knowledge does not travel easily, firms tend to initially seek for solutions locally (Gertler, 2003). The social learning process theory suggests that it is easier to share information and knowledge through face-to-face communication among those already sharing similar attributes, such as same languages, culture, community experiences, and knowledge training. In a successful learning region, local knowledge is transmitted frequently among firms that are involved in a similar market so that the sharing of tacit knowledge, e.g., via formal meetings, past interactions, or social networking, will increase the likelihood of finding the right solution. This argument consists with the specialization argument that firms share similar knowledge domains are more likely to benefit from each other.

In summary, knowledge spillovers have been the important external effects in the agglomeration economy yet are a very abstract concept. In particular, knowledge spillovers mean the knowledge flows among firms. There are three important knowledge transfers mechanisms (or knowledge flows) that need to be known to understand the process of knowledge spillovers: 1) industrial specification and diversification (Glaeser et al., 1992), 2) labor mobility (Almeida and Kogut, 1999; Boschma and Frenken, 2009; Rosenkopf and Almeida, 2003), and 3) local research universities (Breznitz and Anderson, 2005; Feldman, 1994; Youtie and Shapira, 2008). Glaeser's work proposes an opposing argument to Marshall. He suggests that diversification increases the growth of the city, but not the specialization of industry. The second mechanism is via labor mobility that location matters in capturing the circulation of the human capital of key individuals moving in the same region. The third mechanism reveals that the presence of the local universities increases knowledge diffusion from academia to industry by either formal or informal collaborations.

To conclude, previous sections describe different regional resources in the agglomeration economy. The collocation of firms in a region generates the external effects of knowledge spillovers through three different mechanisms, including industrial diversification, labor mobility and the presence of local universities. The previous section discussed the contrasting debates between specification and diversification in the agglomeration economy. The following section focuses on the role of university knowledge and labor mobility on the knowledge flow process.

University resources

One knowledge flow mechanism is the transfer of university knowledge. As Mowery (1998) points out, the US innovation system saw a structural change in the 1980s. With the pressure of urging competitiveness and fewer returns from conducting R&D internally, many large corporations (e.g., AT&T, GE, and Du Pont) in the US began to downsize their R&D operation units. Firms started to adopt a new division of labor approach in the US innovation system by relying more on external knowledge sources. One important change was the increase of university-industry collaborations. Empirical studies also present that firm can access to university knowledge via formal or informal channels. The university-industry collaboration can be operated via several paths, such as consulting, student internship, technology transfer, and being a policy practitioner in an economic and business development program (Rahm et al., 1999). With the Carnegie Mellon Survey, while controlling types of industries, Cohen, Nelson and Walsh (2002) found that the influence of public research (e.g., university knowledge and scientific publication) on industrial R&D is substantial for generating new ideas and is "at least as great as the effect of that originating from rival R&D" (Cohen et al., 2002, p. 21). At the national level, Fernadex-Ribas and Shapira (2009) present that the host country's scientific capacity is important to attract innovative activities by multinational corporations.

In addition, universities are tied closely with regions, particularly within the knowledge-intensive districts. To understand the determinants of the success of Silicon Valley, Saxenian (1996) emphasizes the role of Stanford University as the mediator bridging government laboratory, local entrepreneurs, and small business. The role of universities is more than education and research, but also associated with disseminating and exchanging intellectual discoveries to local organizations. Breznitz and Anderson (2005) highlight that the clustering of the biotechnology industry in the Boston metropolitan area is due to the following reasons: locality, skilled labor force, universities, hospitals, commercial space, and information exchange. Their results suggest that universities are also active participants in the cluster, collaborating with local business. Not only does the university contribute to transferring discoveries to local firms, but in return, they also benefit from access to research practices and funding from local firms. However, transferring technologies from university to industry is not an easy task. It requires repetitive communication and trials to transform basic knowledge into a commercial reality that is satisfactory from an industry perspective (Schimank, 1988). Therefore, collaborating with university is likely to risk in low chance of commercialization.

Zucker et al. (2002) find that collaboration between U.S. star scientists and firms are geographically bound. Although university researchers conform to the norm when publishing their discoveries, those scientific journal papers are difficult to comprehend without direct instructions. For example, scientific papers often simplify the details of the experiments. Feldman (1994) found that the research capacity of universities in a region greatly benefits overall innovation activities in a region. Universities are one source of generating start-ups for the local economy while transferring technologies from scientific research to commercial use, either via university professors or via university-industry collaborations. Hence, the higher percentage of innovative small firms in a region, the more likely it is that university knowledge will be useful and accessible by local business. A bigger pool of innovative small firms will also increase the demand for external knowledge seeking. University professors are more likely to develop the necessary skills to collaborate with local business. To examine this argument, Hypothesis 2a tests whether the net effect of regional ecology is explained by the research capability of local universities. Particularly, it will test if the presence of many small innovative firms explains the increasing value of local universities. Meanwhile, it will also test whether the direct effect of local university knowledge is positively associated with regional innovation performance.

Hypothesis 2a: Regional university knowledge mediates the effect of the SME-dominated ecology on firm's probability to commercialization.

Labor mobility

Another knowledge flow mechanism for firms concerns labor mobility. For firms, skilled engineers and developers own valuable knowledge related to the core tasks in an organization. Economists see individual mobility as a dynamic event representing the exchange of resources, especially information and critical knowledge, among firms and regions. Hiring mobile workers from neighboring firms is an efficient way to earn external knowledge from other firms (Breschi and Lissoni, 2009; Oettl and Agrawal, 2008; Singh and Agrawal, 2011; Stolpe, 2001). According to Saxenian (1994), the success of Silicon Valley has been contributed to by the unique culture of decentralized organizational structures and the high-velocity of labor turnover in the region.

Jaffe (1993) also notices the positive relationship between mobility and inter-firm knowledge flow, suggesting that this relationship is geographically constrained. These moving workers might keep their previous ties with old colleagues in neighboring firms, which can increase the intra-regional knowledge flow among firms (Jaffe et al., 1993). Jaffe is the pioneer researcher who views patent citation as a paper footprint of knowledge flows. Following Jaffe and his colleagues' method, later empirical studies show the positive relationship between mobility and localization of knowledge flow in the semiconductor industry. The knowledge footprint of mobile engineers, either with prior colleagues or with new colleagues is bound by geography (Almeida and Kogut, 1999). Similar results are presented by Agrawal et al. (2006) using US data, Lenzi (2009) in Italian data and Song et al. (2003) focusing on Taiwanese patents. However, whether the encouragement of labor mobility is always good for cluster sustainability is still an intriguing open question.

The hiring of mobile workers is positively associated with overall knowledge learning from sourcing firms to destination firms (Singh and Agrawal, 2011). Singh and Agrawal argue that by recruiting new workers, firms can extend their search space for knowledge to a broader area. Prior empirical studies have tested inventor mobility in innovative productivity by measuring: 1) the quantity of patents (i.e. the number of post-move patents), and 2) the quality of patents (i.e. the number of forward citations) produced by mobile inventors. Taking into account the possible simultaneous causality issue between mobility and innovative productivity, Hoisl (2007, 2009) still found that movers generate more patent application than non-movers do. By the same token, Trajtenberg (2004) observes that mobile inventors are having more domain-specific (i.e. more concentrated in technological fields) and valuable (i.e. more cited) patents.

Mobile inventors can innovatively outperform non-mobile inventors because they are not yet accustomed to the working practice of the hiring firm, thus are likely to come up with new ideas and serendipitous good results. Similarly, mobile inventors are important for decoding both tacit and codified knowledge. For example, they better could know how to decode external information using different approaches compared to their non-moving peers. Even codified documents (e.g. patent disclosure; someone else's programming codes) contain some level of un-codified information that requires extra explanation. One example is the Bessemer steel process invented by Henry Bessemer. In 1855, Bessemer sold his patent of the new steel making process to several large ironmasters. However, his purchasers could not get the process to work. Bessemer ended up starting his own steel company (Gordon, 1984).

Mobile workers are also likely to succeed in the innovation process at a later stage of innovation development. The argument is that they are better at combining technologies and have a stronger chances of commercialization due to the heterogeneous skill sets they possess (Fleming, 2001; Singh and Agrawal, 2011). In contrast, mobile workers might also perform worse since they are new to the firm and lack the market knowledge and capacities of the new firm.

Labor mobility could increase the density of social network not only for individuals, but also for communities and firms within a region. One reason is that workers tend to move locally. Casper (2007) studied the growth of the biotechnology industry in San Diego by observing the formation of career affiliation networks among a pool of senior managers over time. Casper shows that most managers developed social ties through job-hopping, which also indirectly contributed to the whole biotech network in the San Diego area. This network became sustainable through shared career experiences, which further increased the influx of spin-offs, skilled labors, and new innovative ideas in the region. By contrast, a sparse network can lead to the decline of a region. A recent study presents that the creation of a well-connected network among local firms is crucial for the region to keep developing (Breznitz and Taylor, 2009). Using the firm network data in Atlanta, Breznitz and Taylor found that the lack of connected social networks has caused many entrepreneurs and venture capital leaving this region.

In summary, labor mobility is an effective knowledge flow mechanism among firms in the same region. Active labor mobility in a region is likely to increase the innovation activities of firms, as well as regional aggregated innovation growth. It can be argued that the positive effect of a small firm dominated region on innovation performance comes from an increase in labor mobility in small firm dominated regions for the following reasons. Firstly, the concentration of small firms encourages the occurrence of entrepreneurial activities. Secondly, the Silicon Valley case study describes a network-like structure and a very encouraging culture for job change among firms (Saxenian, 1994). In contrast, Route 128 is a region dominated by large firms where the culture is more conservative towards job-hopping from large firms to small firms. As a result, Saxenian's findings show that the Silicon Valley area is comparatively successful than the Route 128 area in the electronics fields. By introducing the regional ecology concept, this study argues that the positive loop of mobility in Silicon Valley is not a cultural reason, but is an ecological reason. Hence, this study predicts that inter-firm mobility mediates the knowledge flow process of the regional ecology. The net effect of the presence of many small firms that dominate a region should drop once we add the labor mobility process to the model.

Hypothesis 2b: Regional inventor mobility mediates the relationship between regional ecology and firm's probability to commercialize their invented technologies. The effect of small firm ecology is due to the increase of mobility *rates in a SME dominated ecology, with high mobility rates increasing the commercialization rates.*

1.2.6 Effects of regional ecology on innovation performance by firm size

Based on existing literature, regional recourses are finite. According to the resource-partitioning model, the resources an individual firm obtains are contingent on the organizational ecology in which the firm is located. Based on the first hypothesis that small firm dominated ecology increases regional innovation performance, the following section further develops theoretical arguments regarding the interplay of firm size and types of collocation on innovation performance. Whether small firms benefit from being concentrated is one fundamental question following Marshall's theory.

The Marshallian thesis: Small firms benefit in the SME dominated ecology

According to Markusen (1996), in the Marshallian district, the region agglomerates a large percentage of small firms. The current research refers to this type of region as the small firm dominated ecology. Firms collocated in this district/region possess competitive advantages regarding their flexibility and specialization (Pyke and Sengenberger, 1992). The collective efficiency allows specialized small and medium size firms to catch up with the technology used by large firms.

Collocation is more important for small to medium-sized enterprises (SMEs) (Baumol et al., 2007; MacKinnon et al., 2002) for two reasons. First, small firms agglomerate to share infrastructures and they need to use local proximity as an advantage to minimize transaction costs in the constantly innovating economy (Simmie, 2005). Secondly, a small firm dominated ecology increases the likelihood of informal communication to share ideas and technological knowledge within the local network (Casper, 2007). In order to obtain successful commercialization, small firms need to acquire the right information to better target the market demand. Therefore, small firms benefit more from the presence of the specialized business services because small firms need to access to information and resources that are complimentary to their services (Feldman, 1994).

In addition, small and medium size firms are largely dependent on the resources they can mobilize locally (Crevoisier, 2009), such as local knowledge and local customers. In contrast, large firms often have greater R&D resources (e.g., larger R&D budget and teams) (Cohen and Levinthal, 1990) and multiple branches in different locations. Large corporations can also send brilliant scientists from one city to another more easily (Zucker and Darby, 1996). Put differently, large firms are less constraint by the spatial boundary for external knowledge resources. This study proposes that small firms are more sensitive to the positive externalities from a SMEs dominated ecology than large firms are, arguing that small firms learn from each other, leading to higher rates of innovation (commercialization) in the presence of many small firm inventions. The following hypothesis summarizes this prediction.

Hypothesis 3a: As the percentage of small patents increases (toward SME-dominated ecology), firms increase the likelihood of commercializing their patented inventions. (to test the direct effect of ecology)

Hypothesis 3b: The difference in commercialization propensity between large and small firms is larger in a SME dominated ecology than in a large-firm dominated ecology. (to test the interaction effect) [1].

The moderation effects between firm size and access to external knowledge flows

Cohen, Nelson and Walsh (2002) indicate that public research is helpful for suggesting new R&D ideas and completing existing R&D projects. They also suggest that large firms benefit disproportionately more than small and medium size firms in appropriating public research. The concentration of large firm innovation can increase the value of the university knowledge. Agrawal and Cockburn (2003) find that the presence of a large hub (an anchor tenant firm) in a metropolitan area generates positive regional externalities for the local innovation system (particularly SMEs) by making university knowledge more likely to be absorbed.

The contrasting argument is that the concentration of small firms increases the value of university R&D because it forces firms in the small firm dominated region to search for affordable external knowledge. For small firms, internal R&D could be too

1 This dissertation recognizes the contrasting theories predicting the effect of regional ecology on commercialization by firm size. One is the Anchor-tenant thesis. It says that large corporations bring positive effects to the local economy by spinning off new firms and increasing in-flows of related firms and skilled laborers. This model was particular popular during 1930s to 1970s, for example the Big Three auto corporations in Detroit, 3M in Minneapolis, and Boeing in Seattle, where small firms positioned themselves being specialists nurtured by large generalists (Markusen, 1996). Empirically, Agrawal and Cockburn (2003) present that the presence of "Anchor Tenant" (large and R&D intensive) firms generates positive regional externalities for the local innovation system (e.g., SMEs) by making university information more likely to be absorbed. The other alternative assumption is the power dynamics thesis. There are concerns of the differential influences of geographic factors on the capacity of firms to learn and innovate. As Boschma (2005) once question, the advantages of collocation of firms is taken for granted as if all firms in the region can access to shared resources with equal chances, and as if all firms in the region are willing to add resources to the pool and encourage sharing. Florida and Kenney demonstrate that for some US firms, even when agglomerating, do not reap the advantages of geographic proximity expected from the industrial district paradigm. Small firms could suffer from the dominance of the anchor tenant due to power differentials in the region (Kenney & Florida, 1994).

costly and they are more reliant on local resources, implying that university R&D could be a very good substitute. Therefore, this study argues that regional ecology moderates the effect of university at the regional and firm levels.

The second issue is that if regional ecology increases the value of local research universities, then does positive impact of university R&D vary by firm size. One argument is that small firms benefit more from universities. For instance, in Japan, small firms are more productive than large firms while collaborating with local university. The projects of small firms are likely to have concrete goals and the employees of such firm have greater decision making autonomy when collaborating with university professors (Motohashi, 2005).

By contrast, some scholars have argued that large firms benefit from local research universities more than small firms. Projects undertaken by large firms tend to have a long-term goal, which increases the degree of uncertainty and time. In addition, university research is usually more basic-oriented, which requires more years to turn an R&D discovery into an innovation. Hence, we can presume that larger businesses are more likely to capture or finance new technology. I predict that university R&D is likely to moderate the effect of firm size.

Hypothesis 4a: The probability to commercialize large firms' patented invention increases more than that of small firms' patented invention with increasing university R&D expenditure in the region (i.e., Large firms are benefiting more from being surrounded by research universities than SMEs).

Similarly, empirical studies suggest that the advantage of hiring mobile engineers is to combine technological distant knowledge (Song et al., 2001, 2003). However, one of the learning traps in large firms is the tendency of staying in the fields they are more competent (Levinthal and March, 1988), making them less likely to adopt ideas from new hires. This view is corroborated in a recent empirical study in Japan that large firms cite more of their own patents (rather than other firms) over small firms (Motohashi and Muramatsu, 2012). Therefore, we predict that large firms receive less labor mobility benefits than small firms.

Hypothesis 4b: The probability to commercialize small firms' patented inventions increases more than that of large firms' patent inventions with increasing regional mobility (i.e., SMEs benefits more from labor mobility than their larger counterparts).

1.2.7 Summary

Innovation is the engine for economic growth, for both regions and firms. Regions are not only locations, but also an organic entity. Innovation takes place in these spatial entities that provides external resource. According to the literature, there is a lack of dialogue between organizational theorists and economic geographers on the topic of regional innovation and firms. This reveals a concern the tension may exist between regions and individual firms. The worry is that the goal for achieving an innovative region could differ from the goals of individual firms participating in the innovation market. To understand the role of geographic proximity as a platform of knowledge flow to enhance the process of innovation in firms (Feldman, 1996), it is important to re-conceptualize geographic proximity as "regional ecology" to emphasize the organizational ecology of the region, rather than merely considering the distance proximity of firms.

The goal of this research is to improve the contentious theoretical concepts in current theory on the regional innovation system. Therefore, we will redefine the concept of agglomeration by decomposing the size concentration of firms in a cluster. The thesis contributes to existing literature and to the ecological understanding of innovation performance by clarifying the following research agenda:

1) The role regional resources play in explaining innovation.
2) The role regional ecology plays in enhancing or decreasing innovation.
3) The differential effects of regional contexts vary by firm size.

Based on the literature review and hypotheses developed in this chapter, Figure 1.1 presents a conceptual model of this dissertation. The dependent variable in the model is the innovation performance that is operationalized as 1) the patent per capita (inventions) and 2) the propensity to commercialize a patented invention (commercial innovations). To test the regional innovation system theory, the independent variables include both regional and firm level factors. The regional level factor emphasized in this research is the size concentration of firms in a region, framed as “regional ecology.” Other regional level resources representing the knowledge flow mechanism are regional diversification across industries, the university knowledge, and labor mobility. I develop my hypotheses following the Marshallian tradition, aiming to investigate how vital is the small firm dominated ecology. Figure 1.1 is a diagram presenting the overall conceptual model tested in this research.

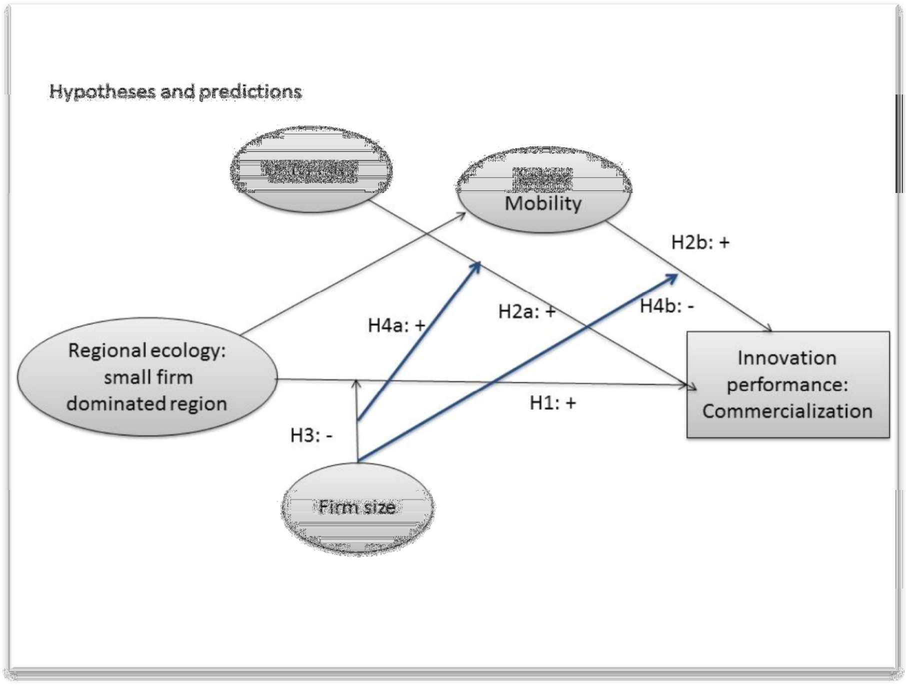

Figure 1.1 Conceptual model and predicted hypotheses

First, we propose testing the influence of regional ecology on innovation performance. The first hypothesis is that a small firm dominated region is expected to have a positive impact on innovation performance (patenting activity and commercialization) (H1a and H1b). Firm size will be used as a control when testing the net effect of regional ecology. As Figure 1.1 depicts, two mechanisms of information flows, university knowledge and labor mobility, are also examined in this research. The net effect of university knowledge is positive (H2a) to the innovation performance. Additionally, the net effect of mobility (H2b) is expected to be positive to commercialization.

Figure 1.1 also illustrates the conceptual model for answering the third research question, "how does the effect of regional ecology vary by firm size?" Firm size is predicted to have a moderating effect on the influence of regional ecology on innovation performance. According to literature, the line of H3 represents the moderation effect of firm size and regional ecologies. I propose the Marshallian thesis, suggesting that the interaction effect of small firm dominated ecology and firm size is positive for small firms.

Furthermore, this study predicts that university knowledge could be more useful for large firms than small firms. Hence, the moderating effect of university on the influence of firm size is positive (H4a). In addition, labor mobility is more useful for small firms, hence I predict the moderating effect of mobility on the influence of firm size will be negative (H4b). The next chapter describes the methodology, data sources, and key measures in more detail.

CHAPTER 2

METHODOLOGY, DATA, AND MEASURES

To examine the afore-mentioned research questions and hypotheses, I need detailed information about innovation activities at the regional level and the firm level. The data should be able to describe the process of innovation together with firm level characteristics and environmental characteristics.

Estimates are based on a novel data consist of multiple data sources, including an US inventor survey and several archival datasets (e.g., USPTO online database, and PATSTAT). The major data is the "The Georgia Tech/RIETI 2007 Inventor Survey: Inventors and Their Inventions[2]" (The GT/RIETI survey). The survey was administrated in cooperation with the Research Institute of Economy, Trade and Industry of Japan (RIETI) between June and November 2007. The next section introduces the GT/RIETI survey regarding its sampling design, survey instruments, and variables used for this study.

2.1 The GT/RIETI survey

The sample of The GT/RIETI survey was from the granted United States (US) patents filed between 2000 and 2003 (in terms of the first priority date). Those patents were included in the Organization for Economic Co-operation and Development (OECD) triadic patent family. Triadic patents are patents filed in both the Japanese Patent Office

---

[2] As one of the research members, I participated intensively in administrating the survey. I was involved in developing the survey, including modifying the questionnaire, programming the web survey, and managing the dataset.

(JPO) and the European Patent Office (EPO), and granted in the United States Patent and Trademark Office (USPTO). This suggests that those patents are globally focused. We randomly selected 9,060 triadic patents stratified by National Business Economic Research (NBER) technology class (Hall, Jaffe, and Trajtenberg, 2001). For each patent, we selected the first US inventor, and we collected US addresses of the US inventors from the EPO database and other supplementary sources (e.g., phone directories). If no valid address was available, we took the next US inventor on the patent. After randomly drawing one patent for inventors holding multiple patents, we mailed out the questionnaires to 7,933 unique inventors. We did not send multiple surveys to the same inventor because doing that would probably increase the non-response rate. The number of patents belongs to each unique inventor was coded as the sampling weight (inverse probability of selection) to adjust for multiple-patent inventors. For descriptive statistics, we used sampling weight to better estimate the expected value of the measures. Table 2.1 shows the distribution of patents per inventor in the survey sample. About 10% of the sample includes inventors who show up more than once in the sample. This also suggests that there are very few continuous inventors in the sample. Particularly, during our four-year window (among USPTO patents filed during 2000-2004), 95% of small firm inventors only patented once, suggesting that large firms' inventors are having more continuous outputs than that of small firms.

Table 2.1 Number of patents per inventor in the sample

Number of patents per inventor	Full sample		Small entity	
	Frequency	Percent	Frequency	Percent
1	7,124	89.8	1,110	94.7
2	624	7.9	48	4.2
3	115	1.4	10	0.8
4	43	0.5	3	0.3
5	13	0.2	0	0
6	10	0.1	0	0
7	4	0.1	0	0
Total	7,933	100.0	1,171	100.0

The survey was designed in mixed-modes, including both web and mail survey. We sent out questionnaires and cover letters (included information of the survey URL) to 7,933 unique inventors. They could respond either by post-mail or by web. In between the two waves of mail-out packages, we sent a reminder (the thank you note) to all inventors in the sample. We received 1,919 respondents, yielding to a 24% response rate and a 32% adjusted response rate by eliminating undelivered cases. Of the 1,919 respondents, the percentage of mail and web is 63% and 37% respectively. We ran tests for non-response bias and survey-mode bias to avoid self-selection problems in the data. The test results did not show much significant difference between response vs. non-response and web vs. mail groups. Alternatively, we did not see significant differences in patent-related measures, such as the number of references, the number of inventors, and the number of technological classes, while doing the comparison. However, we did find that web respondents are younger than mail respondents are, as well as receiving more forward citations than mail respondents are. This result suggests that a mixed-modes strategy facilitates a better coverage of the sample. Of the total valid respondents, we have 1,806 inventors affiliated with firms. In the survey, inventors from large firms (employee > 500) account for 81%, mid-size firms (100 < employee <= 500) for 7.7%, and very small firms (employee <= 100) for 11.2%. Next, the following section introduces the research design for collecting regional level data.

2.2 Spatial data of innovation activity in US metropolitan areas

The second database includes the location of the US inventors on all patents granted by USPTO from the 2000-2003 cohorts (total N = 341,915). I geocoded the first inventor with a US address for each utility patent with the help of ArcGis software based on the boundary files of metropolitan statistical areas (MSAs) and the 2000 ZIP Code Tabulation Areas (ZCTA) data[3]. Of the 341,915 patents, 271,113 patents (79%) included valid zip code information and were successfully geocoded on the map (see Figure 2.2). After joining with the consolidated metropolitan statistical area (CMSA) relation table and excluding patents by university and non-profit organization, this yields us 199,507 patents correspond to 279 MSAs. The average number of patent grants from 2000 to 2003 in a region is 745, with a min of one, and a max of 25,185.

Throughout this study, we choose metropolitan Statistical Areas (MSA) as the approximation of the "region," Christopherson and Clark (2007) had discussed the legitimation of using MSAs as adequate proxies for regions. I mapped those USPTO patents based on inventors' addresses rather than assignees' locations so that this

[3] These boundary files are defined by the US Census Bureau 2000 data. Sources are available at: http://www.census.gov/geo/www/cob/ma_metadata.html#msa

geographies (regions) in this paper represent the commuting spaces of the labor market instead of headquarters of firms.

For regression analyses, we recognize the heterogeneity across industries in terms of knowledge domain, business strategy, and so on. Figure 2.4 to Figure 2.9 (p.76-p.78) demonstrate the mix of large and small firm concentration that varies across technological fields in a region. For example, in the chemical field, the proportion of small firm patent in New York MSA (13%) is lower than the national average, but in Los Angeles MSA, the proportion of small firm patent in chemical field (32%) is higher than the national average. To address the heterogeneous firm concentrations by fields, I measure regional variables (regional innovation performance and regional ecology) using region-technology pairs. This is to calculate the MSA-technology level measures, instead of the overall average of the MSA level measures. Supplementary regional data comes from archival dataset (sources are like National Science Foundation and the Census Bureau).

2.3 Key measures

This section introduces key variables used for analysis to test the impact of regional ecologies on innovation performance. I begin with describing dependent variables, including firm level innovation performance and regional level innovation performance. Then I introduce key independent variables, i.e., regional ecology, mechanisms of knowledge flow, and firm size. I also describe control variables used for analyses as well as alternative explanations. Table 2.3 presents the list of variables used in this study and the data sources.

2.3.1 Dependent variables

One common measure of innovation activity is the R&D expenditure of a private organization. Sometimes the measure is counts of patents (Jaffe, 1998). Other times it is counts of innovations (Acs, Audretsch, and Feldman, 2002), which is often seen as a more direct measure. Empirical studies have shown that patents provide a reliable measure of innovative activities (Acs, Anselin, & Varga, 2002) because it is similar to regression results using counts of innovations as the dependent variable at the metropolitan area level. Compared with bibliometrics patent data, the survey of inventors on patents provides detailed information of the innovation process. The advantage of using survey data is to obtain a more in-depth understanding of the R&D process and the use of the invented technology at the time inventors were involved in that certain patent project.

In this study, I combine these two approaches by using a unique survey of US patent inventors. In this case, patent is the proxy for new technological invention. Innovative activity is then measured as the commercial use of the patented invention. The emphasis on the use of patent (i.e., commercialization) was less explored in past literature. Hence, the major dependent variable is the commercialization rate of a patented invention.

DV1: Commercialization at the project level

In the GT/REITI survey, we asked respondents a series of questions whether the patented invention was commercially used, including if the patent is 1) commercialized in a product/process/service by the applicant/owner, 2) licensed by (one-of) the patentholder(s) to an independent party, 3) established a start-up firm by the respondent or any of respondent's co-inventors. If any of the above questions are checked "yes", then the commercialization variable is coded as 1, otherwise it is coded as 0.

Out of 1,742 complete cases (including university inventors), 971 (56%) respondents reported that their patents are used for commercial purposes. I exclude university inventors when conducting analyses in this study. The average time gap between the filed date and the launched date of a patented invention is 2.4 years. . In the sample, most of the patents filed between 2000 and 2003. The commercialization rate in 2000 is 55%, following by 56% in 2001, 55% in 2002, and 49% in 2003, which shows a level trend. Descriptively, small firms with less than 100 employees are more likely to use their patent inventions for commercialization (75%) than large (51%) and medium-sized (60%) firms are.

DV2: Regional innovation performance

To measure innovation performance for each metropolitan, two variables are constructed. One is the patenting activity in each MSA, as a proxy for a MSA's capacity for innovation. We counted the number of granted USPTO patents filed between 2000 and 2003 for each region (MSAs) and divided by the 1,000 population based on the population data of Census 2000. We call this variable "rate of innovative activity", indicating the patenting activity per thousand populations for each region (MSAs)[4]. The second variable is the commercialization rate per MSA. Using the GT/REITI survey, I

[4] To eliminate potential ratio variable problems, I later use the total count of patents in each MSA as an alternative measure of the regional patenting activities.

calculated the regional commercialization rate per MSA (percent of triadic patents in a region that are commercialized). The aggregated mean of the commercialization dummy is calculated based on respondents' location if in the same metropolitan statistical areas.

2.3.2 Key explanatory variables

The following sections describe key independent variables used in the study.

(1) Regional-level independent variables

Regional ecology

The main explanatory variable is a measure of the mix of firm size in a region. To calculate this measure, I collect the population patents of the 2000-2003 cohorts filed in USPTO. Then, I coded each patent as small business, or university/government lab patents based on the USPTO patent fee maintenance database, which includes a field designating patents as belonging to "small entities" (defining as independent inventor, a small business [generally less than 500 employees from manufacturing], or a nonprofit organization [e.g., university]). While aggregating the full population to the MSA level, we create a variable, *SMEpat*, measuring the number of patents produced by a small entity in the MSA-technology, excluding those universities and non-profit organizations. *PctSMFPat*[5] is measured as the percentage of small firm patents in MSA i and technology field j.

---

[5] Throughout the list of patents filed between 2000 and 2003, I manually identified university and colleges based on the assignee field on patent documents. The *PctSMFPat* measure excluded university patents so that the share of patents owned by small entities is mainly the ratio of small private firms out of industrial patents granted in a MSA-technology.

$PctSMFPat_{ij}$ = $SMEpat_{ij}$ / (# of total industrial patents)$_{ij}$ *100

Diversity of technology field

I measure the concentration of innovative activities across MSAs on the dimensions across technological fields. One simple measure is the density measure (Carroll, 1985), the number of patent assignees in a MSA. The other measure derives from the Herfindahl index to characterize the degree of diversity. This measure is to characterize the distribution of technological fields in a particular MSA (Agrawal et al., 2010). This measure is similar to the "generality" and "originality" measure referring to the basicness of patents developed by Hall and her colleagues (2002). Whereas Hall et al. calculated the concentration of citations of a patent, I measure the concentration of firms participating across technological fields in a MSA. I calculate the inverse measure to represent the normalized diversity of technology field in a region, called $Tech\_diversity_{msa}$. The formula is as

$Tech\_diversity_{msa} = [1 - \sum_{nber \in 6} (\frac{N_{msa}, nber}{N_{msa}})^2] \frac{N_{msa}}{N_{msa}-1}$, where nber is the set of six NBER technology classes in which the MSAs were issued with more than one patent. This standardized measure is between zero and one. The larger the number means that the MSA is more diverse across technology fields. As the number reaches zero, it means that the MSA is getting more concentrated. In other words, this measure is the opposite of the concentration measure.

The amount of university R&D expenditure in a MSA

In addition to testing the ecology effect on innovation generally, we also test the role of university as the knowledge channel for innovation in a region. We collected R&D expenditure data of universities and colleges for each metropolitan area based on the report of National Science Foundation (NSF) - Science and Engineering Indicator published in 2002.

The mobility rate in a MSA

At the regional level, the labor mobility rate for each metropolitan area was collected from the USPTO patent archival database, based on the population patents from the 2000-2003 cohorts filed at USPTO. This study chooses to examine inventor mobility using the US patent database for several reasons. First, patents are public accessed documents, which make the trace of R&D outputs of inventors explicit. Relating to the first reason, each patent lists the information of the hometown of the inventor, thus researchers can use this information to identify inventor's region of residence. Third, patent data represents a sample of high skilled workers, which allows us to emphasize specific types of labor mobility, instead of general labor mobility[6]. Based on Lai, D'Amour and Fleming's (2009) Inventor Dataset, a longitudinal patent inventor database from 1975 through 2006, I construct a variable measuring the inventor mobility rate for each metropolitan statistical area.

I first collect a subset of the data of USPTO patents filed during 2000 and 2003 (patents = 780,981, inventors = 134,823). I exclude inventors with only one patent. For

[6] For this reason, I did not choose the general residence migration rate from the data of Current Population Survey (CPS) because the CPS is not able to identify the situation of job turnovers, particularly in the high-tech sector.

each inventors with multiple patents filed during that period, I wrote a simple syntax to detect if the inventor had a move between assignees. To detect possible typos in the assignees' names, I combined the Soundex and the Compged algorithm since these two algorithms are complementary to each other. I found that 29% of inventors had moved. This percentage of movers in USPTO patents is similar to the results in the RIETI/GT survey, which reported approximate 26% of mobile inventors. Then, to calculate regional mobility rates, I compute the mean of moving events by using the USPTO patent inventors data aggregated at the regional (MSA) level.

(2) Firm-level independent variables

Firm size

In the RIETI/GT inventor survey, we asked respondents to report types of affiliations (e.g., private firms, university, government laboratory, and other organizations) they worked with at the time of the project. The question categorized firm size into four employment-size categories based on the number of employees (over 500, 251 – 500, 100 – 250, and less than 100). Of all respondents working for private firms, there are 113 cases with missing value about the size of the firm. To reduce the missing cases, we collected supplementary data (e.g., company websites, and USPTO patent fee maintenance database) to help assess the size of the respondents' organization affiliations. This yields 1,849 valid respondents who answered the firm size question. The unweight share of large firm (> 500 employees) accounts for 80% of respondents. The share of very small firms (< 100 employees) is 12% in the sample and 7% are the medium sized firms. The weighted statistics are similar.

In this study, I construct a dichotomous variable of firm size by defining large firm, coded 1, as a firm with more than 501 employees, and small and medium size firm, coded 0, as a firm with less than 500 employees. I choose to use 500 employees as the cutoff point for defining large firms and SME defined by the U.S. Small Business Administration standard.

Table 2.2 Distribution of firm size of the respondents in the GT/REITI Survey

Firm size	Unweight		Weighted	
	N	%	N	%
A firm with more than 500 employees	1486	80.4	1786	81.8
A firm with 251 to 500 employees	76	4.1	79	3.6
A firm with 101 to 250 employees	63	3.4	70	3.2
A firm with less than 100 employees	224	12.1	247	11.3
Total	1849	100	2182	100

2.3.3 Control variables: alternative explanation

This study also takes into account alternative explanations to predict the commercialization of a patented invention. First, according to Haltiwanger et al. (2010), young and startup firms contribute substantially to regional job creation and growth. Hence, I control the percentage of young firms (less than 5 years) in a region (MSA-level) aggregated from the GT inventor survey. At the project level, I control the percentage of the inventors' time spent on basic research because projects involving greater basic research are less likely to be commercialized than applied research. Based on the quadrant framework of Stokes (1997), the more basic-oriented research suggests higher cost and more uncertainty compared with need-driven research aiming to answer existing answers. I also control the size of a project by the total number of man-month spent on a patent project. Technologies came from larger projects are more likely to be commercialized than smaller projects. The explanation is that a bigger project should be more likely to generate at least one commercializable invention than a smaller project.

Another important factor for innovation performance in the innovation literature is the formal and informal collaborations among firms. A collaborative project among multiple organizations increases knowledge sharing and thus increases the likelihood of commercialization. In addition, Rothaermel and Deeds (2004) also found that the alliances of firms increase the likelihood of developing marketable technologies. Hence, I control the effect of business alliance or informal collaboration if occurred in the focal patent project. In the RIETI/GT inventor survey, we asked inventors to indicate if they have collaborated with others either formally or informally for the focal patent. I construct a dummy variable for any collaborator by coding it as 1 if there were any collaborators (formal or informal) on the focal patent, and 0 otherwise.

Furthermore, technologies with higher technological value are more likely to be commercialized. This study controls the technological value of the invention using a self-assessment measure based on a recent PATVAL survey of Gambardella et al., 2008. We asked our respondents to rank their invention in a four-point scale (i.e. top 10%, 10%-25%, 25%-50%, and bottom half) compared with other technologies invented in the US at the same time. Invention collaboration among firms indicates an important mechanism of inter-firm knowledge sharing and a higher probability taking the patented invention to the second stage of innovation. Finally, I also control the number of inventors on the patent, the issued year and technology fields based on the National Bureau of Economic Research (NBER) classification (Hall et al., 2001).

Table 2.3 List of variables and brief descriptions

Variable		Description	Data source
Regional characteristics			
MSA-field	Rate of commercialization	Percentage of commercialized patents	GT/RIETI Survey
	Rate of inventive activity	USPTO patents per capita	USPTO
	Regional ecology	Percentage of small firm patents	USPTO
MSA	Number of firms	Number of patenting assignees in a MSA (take the logarithm)	USPTO
	Diversity of technology fields	A measure of diversity of technology fields in a MSA	USPTO
	MSA mobility rate	Inventor mobility rates in a MSA	USPTO/Lai et al.
	Academic R&D expenditures	Amount of university R&D expenditure in FY 2002 in a MSA (take the logarithm)	NSF
	Rate of startups	Percentage of young firms (less than five years) in a MSA	GT/RIETI Survey
Project characteristics (the unit is patent)			
Large firm (Y/N)		A dummy variable coded as 1 if firm size is more than 501 employees	GT/RIETI Survey
Scale of the project		Inventor-months for the project leading to the patented invention	GT/RIETI Survey
Proportion of basic R&D (%)		Percentage of inventor's time spent in basic research	GT/RIETI Survey
Top 10% Technological value (Y/N)		A dummy variable coded as 1 if the patented invention ranked as the top10% (self-reported)	GT/RIETI Survey
Number of inventors		Number of inventors on the US patent	PATSTAT
Project collaboration (Y/N)		A dummy variable coded as 1 if the research leading to the focal patent had any (formal or informal) collaboration partners	GT/RIETI Survey
Technological field (NBER class)		Dummies of six technology fields are included: Chemical, Computer&communication, Drug&medical, Electrical&electronic, Mechanical, and Others. The reference group is the "others"	PATSTAT
Any venture capital funding (Y/N)		A dummy variable coded as 1 if the patent project had any venture capital funding	GT/RIETI Survey
Patent issued year		The year that the patent was granted	PATSTAT

2.4 Descriptive statistics

Regional variables

To begin, I first plot all the USPTO patents filed from 2000 to 2003 (see Figure 2.2). It is clear that inventions were clustered in the metropolitan areas, particularly those on the east and the west coast. While the full sample covers 279 metropolitan areas, many MSAs had too few patent samples to generate meaningful results; hence I exclude regions with less than 20 patent applications during the 4-year window. Figure 2.3 presents a map of small firm patents ratio in scales across MSAs in the US. The bigger circle means a higher percentage of small firm patents in a region, and vice versa. We see a variant of percentage of small firm patents at the MSA level regardless the differences of technological fields. However, for regions with diverse industries, one concern is that we should compare region-technology pairs to reduce the unobserved bias that firms are only competing with firms in the same industry. We notice there are within-region variations across MSA-technology pairs as shown in Figure 2.4 to Figure 2.9. By doing so, I compare for example NY MSA Chemical with Atlanta MSA Chemical. For example, in the chemical technology field, Cincinnati and Pittsburgh are having similar patent counts, but Cincinnati has only 3% of small firm patents and Pittsburgh has 8% of small firm patents. Therefore, we test our hypotheses at the MSA-technology level. For analysis, I exclude MSA-fields that have less than 20 patents, and this gives us 326 MSA-technology for analyses.

Table 2.4 shows descriptive statistics for the key regional variables but only includes MSAs with more than 20 patent applications. Of the 326 MSAs, the mean of populations is 2.6 million, implying that this study focuses on medium to large metropolitan statistical areas. Rates of inventor mobility in MSAs range from 4% to 40%, with a mean of 20.4%, a value that is similar with previous studies (Rosenkopt and Almeida, 2003; Stolpe, 2001; Marx, Strumsky and Fleming, 2009; Hoisl, 2007). The average diversity rate among six technological classes is 76%, indicating that the diversity rate is 76%, which is close to the national data in the US. If we use the 37 sub-classes to calculate the diversity index, the correlation coefficient between these two measures is 0.75, suggesting that using the top categories is sufficient to represent the diversity across technology fields in a MSA.

By making the MSA and technology dyad, Table 2.4 also reports that the average MSA-technology has about 130 assignees, 523 patent inventions, and 82 small firm patents. The MSA-technology average *patent_per_assignee* is 4.3, with the minimum of 1.2 and the maximum 36, suggesting some level of variation of dominant situation across MSA-technology. The distribution of these variables is not highly skewed. The median of MSA-technologies has a total of 209 patents and 32 of which are small firm patents, although we do have some mega MSA-technology (New York MSA-Chemical with over 3000 patents and Los Angeles MSA-C&C with over 2700 patents) in the sample during 2000 and 2003.

Table 2.5 reports the correlation table of key regional measures. The results show that percentage of small firm patents is positively correlated with regional diversity of technological fields, supporting the claims that small firms provides more specialized technologies and services. The ratio of small firm patents in a region is also positively associated with regional inventor mobility at the 10% significant level. The correlation coefficient between regional commercialization rates and the percentage of small firm patents is positive, but not statistically significant.

<u>Project-level variables</u>

Most of the patent-level variables came from the GT/RIETI survey. I include the descriptive statistics depicted in Table 2.6. On average, among all industry patents (N=1507), 54% were used for commercialization in any kind of approach (either in-house, licensing, or forming a start-up company). Of all the respondents, 39% of them reported that their patents were used for in-house commercialization, and 11% were used for licensing. About 15% of respondents ranked their patented invention at the top 10% among all the inventions in the US in the same period. This number is slightly higher than 10%, but given that we select triadic patents as the sample, we think this number is acceptable. The average number of inventors per patent is 2.7. Of all industry patents, around 2% were co-assigned, 22% were from a collaborative project with multiple organizations. We see a huge gap between co-assignee percentage and collaboration percentage, indicating that the bibliometrics information from patent documents was not able to illustrate the complete story of industrial collaborations (Nagaoka and Walsh, 2009). On average, 8% of the project tasks involve basic research, with a standard deviation of 20%. The average number of forward citations is 3.2 for the full sample.

Next, we break down the data by firm size. As Table 2.6 shows, small and medium sized firms with less than 500 employees have higher commercialization rate (69%) compared with large firms with more than 500 employees (50%, chi-square = 30.7, p<.0001). Large firms and small firms are not significantly different in conducting internal commercialization (40% vs. 36% for small and medium sized firms, chi-square =

1.22, p = .35). Small firms are more likely to choose licensing as the mean of capitalizing their R&D investments (18% vs. 9% for large firms, chi-square = 15.1, $p<.0001$).

About one-quarter of inventors of small firm ranked their inventions with top 10% quality (27%) which is significantly higher than those inventors of large firms (13%, chi-square = 29.3, $p<.0001$). Small firm patents also have a fewer number of inventors in a project than large firm patents (2.52 vs. 2.78 for large firms, chi-square = 4.4, p = 0.02). As predicted, small firms have more collaborative patents than large firms do (28% vs. 21%, chi-square = 7.3, p = 0.017), which is consistent with the assumption that small firms require more external resources to complete an R&D project. Small firms are less likely to file patents in a major field within their region. When checking number of forward citations received, small firm patents (3.63) are slightly higher than large firm patents (3.15), but not statistically significant (chi-square = 1.72).

Next, I look at the data by breaking down technology classes (see Table 2.7). Mechanical technologies have the highest rate of commercialization (60%), followed by electrics and electronics (58%), while drug and medical technologies have the lowest commercialization rate (43%). Table 2.8 shows the result by 34 sub-categories of technology fields. The findings show a great variation of commercialization rates among technological fields of patents, ranging from 27% for drugs to 72% for electrical devices (for those fields with more than 10 patent samples). As Table 2.9 illustrates, this study includes 79 MSAs that vary by commercialization rates, population growth, and inventor mobility rates. The commercialization rates range from 0% to 100%, with a mean of 55%. For example, Atlanta MSA had 67% of commercialization of the patented inventions, 39% of population increase from 1990 to 2000, and 24% of mobility rate. Cincinnati-Hamilton MSA had 40% of commercialization rate and 8% of inventor mobility.

Table 2.4 Descriptive statistics of regional (MSA) variables

Variable	**N**	**Mean**	**S.D.**	**Min**	**Max**	**P50**
At the MSA level						
Pop2000	326	2650392.9	3793263.1	78153	21199865	1187941
# patent application	326	3006.6	4581.3	49	24348	1319
University R&D expenditure (per thousands)	326	382287.3	519608.5	0	2207844	4229.6
Inventor mobility rates	307	20.42	6.41	4.82	40.0	20.2
Diversity Index	326	0.757	0.067	0.27	0.83	0.77
At the MSA-technology level						
USPTO patent counts	326	522.7	941.2	20	9906	209
Rate of inventing activity	326	0.286	0.485	0.014	5.65	0.16
# Small firm patents	326	82.15	136.3	0	838	32
% of small firm patents	326	19.6	14.2	0	66.8	16.8
Number of assignees	326	130.6	196.9	4	1638	54
Commercialization rate (%)	326	55.1	36.9	0	100	54.2
Patents per assignee	326	4.33	3.74	1.24	36.0	3.1

* Only includes MSA-technology with more than 20 patent applications.

Table 2.5 Correlation table of regional (MSA) variables (N = 326)

	1	2	3	4	5	6	7
1 Commercialization rate	1						
2 Inventive activity rate	-0.088	1					
	(.114)						
3 Pct of small firm patents	0.0538	-0.336*	1				
	(.333)	(<.0001)					
4 Technology-field diversity	0.0449	-0.204*	0.253*	1			
	(.368)	(.0002)	(<.0001)				
5 Log(# assignees)	0.0187	-0.032	0.0423	-0.453*	1		
	(.732)	(.564)	(.446)	(<.0001)			
6 Log(university RD$)	-0.0253	-0.031	0.0928*	0.456*	0.594*	1	
	(.649)	(.572)	(.094)	(<.0001)	(<.0001)		
7 Inventor mobility rate	0.0279	-0.210*	0.0932	0.168*	0.223*	0.097*	1
	(0.626)	(0.0002)	(0.103)	(0.003)	(<.0001)	(0.089)	
8 Pct startup firms	0.076	0.009	0.072	0.131*	0.156*	0.213*	0.067
	(.167)	(.871)	(.196)	(.018)	(.005)	(.0001)	(.241)

* Only includes MSA-technology with more than 20 patent applications. P-value is in the parenthesis. *, P<.05

Table 2.6 Descriptive statistics of patent-level variables by firm size

	Full sample		Large firms (N=1213)		SMEs (N=294)		Chi-square
	Mean	SD	Mean	SD	Mean	SD	
Commercialization (y/n)	0.54	0.54	0.50	0.55	0.69	0.48	30.71***
-In-house commercialization	0.39	0.53	0.40	0.54	0.36	0.50	1.22
-Licensing	0.11	0.34	0.09	0.32	0.18	0.40	15.09***
-Start-ups	0.06	0.26	0.03	0.19	0.21	0.43	125.8***
Top 10% tech significance in the US(%)	0.15	0.39	0.13	0.37	0.27	0.47	29.3***
Coassigned patent (y/n)	0.02	0.15	0.02	0.13	0.04	0.21	6.96**
Number of inventors	2.73	1.98	2.78	2.03	2.52	1.78	4.42*
Any collaborator? (y/n)	0.22	0.45	0.21	0.44	0.28	0.47	7.29 **
Basic-oriented project (%)	8.55	20.05	8.76	20.71	7.58	17.02	0.88
Inventor month	19.46	25.91	19.30	26.43	20.21	23.65	0.30
In the dominant field (y/n)	0.31	0.50	0.32	0.51	0.25	0.45	5.52*
Number of forward citations	3.23	5.98	3.15	5.86	3.63	6.47	1.72

Data source: GT/RIETI Inventor Survey; Weighted by inventor-patents weights; firm only cases; *p<.05, **p<.01, ***p<.0001

Table 2.7 Distribution of commercialization by six main technology fields

NBER top category	Commercialization		
	N	Mean	SD
1. Chemical	345	0.52	0.54
2. Computer and Communication	296	0.51	0.56
3. Drug and Medical	214	0.43	0.53
4. Electrics and Electronics	290	0.58	0.54
5. Mechanical	210	0.60	0.53
6. Other	152	0.62	0.52

Weighted by inventor-patents weights; firm only cases; excluding 37 cases with no assignees listed on the patent documents.

Table 2.8 Distribution of commercialization by 34 sub-technology fields

		Commercialization	
	N	Mean	SD
11 Agriculture,Food,Textiles	5	0.43	0.65
12 Coating Chemical	24	0.65	0.51
13 Gas	6	0.67	0.52
14 Organic Compounds	51	0.35	0.52
15 Resins	70	0.49	0.54
19 Miscellaneous/Chemical	189	0.56	0.53
21 Communications	122	0.49	0.57
22 Computer Hardware	33	0.5	0.56
23 Computer Peripherials	31	0.35	0.57
24 Information Storage	33	0.62	0.59
77 Computer Software	77	0.57	0.53
31 Drugs	76	0.27	0.49
32 Surgery & Med Inst.	93	0.52	0.53
33 Biotechnology	23	0.48	0.53
39 Miscellaneous/Drgs&Med	22	0.54	0.53
41 Electrical Devices	38	0.72	0.48
42 Electrical Lighting	27	0.57	0.59
43 Measuring & Testing	50	0.66	0.49
44 Nuclear & X/rays	34	0.54	0.56
45 Power Systems	68	0.54	0.56
46 Semiconductor Devices	46	0.43	0.53
49 Miscellaneous/Elec	27	0.68	0.53
51 Mat. Proc & Handling	42	0.65	0.52
52 Metal Working	41	0.58	0.54
53 Motors & Engines + Parts	34	0.6	0.54
54 Optics	36	0.5	0.52
55 Transportation	25	0.58	0.56
59 Miscellaneous/Mechanical	32	0.67	0.53
61 Agriculture,Husbandry,Foo	10	0.8	0.42
63 Apparel & Textile	9	0.83	0.46
65 Furniture,House Fixtures	7	0.56	0.61
67 Pipes & Joints	11	0.42	0.54
68 Receptacles	15	0.65	0.53
69 Miscellaneous/Others	95	0.59	0.52

Table 2.9 Commercialization rates, population growth, and mobility rates by MSA

CMSA name	Commercialization (%)	Population growth (1990-2000)	Inventor mobility (%)
Albany-Schenectady-Troy, NY	36	1.64	20.5
Albuquerque, NM	56	20.98	31
Allentown-Bethlehem-Easton, PA	75	7.21	25.5
Appleton-Oshkosh-Neenah, WI	14	13.72	4
Atlanta, GA	67	38.93	24
Austin-San Marcos, TX	50	47.69	22
Boise City, ID	100	46.14	4
Boston--Worcester-Lawrence, MA-NH-ME-CT	65	6.67	31
Buffalo--Niagara Falls, NY	45	-1.61	10.5
Canton--Massillon, OH	50	3.25	17.5
Charleston-North Charleston, SC	50	8.32	11.5
Charlotte-Gastonia-Rock Hill, NC-SC MSA	60	29.02	25.5
Chicago-Gary-Kenosha, IL-IN-WI	50	11.14	22
Cincinnati-Hamilton, OH-KY-IN	40	8.89	8
Cleveland--Akron, OH	66	3.01	27.5
Colorado Springs, CO	75	30.2	49
Columbus, OH	67	14.47	29.5
Corvallis, OR	33	10.37	21
Dallas--Fort Worth, TX	52	29.34	22
Dayton-Springfield, OH	43	-0.07	17
Denver-Boulder-Greeley, CO	65	30.37	23.5
Detroit-Ann Arbor-Flint, MI	60	5.19	22
Elmira, NY	60	-4.33	16.5
Evansville-Henderson, IN-KY	100	6.17	9
Florence, SC	86	9.98	39
Fort Collins--Loveland, CO	43	35.11	19.5
Grand Rapids-Muskegon-Holland, MI	63	16.06	21.5
Greenville-Spartanburg-Anderson, SC	85	15.88	24
Harrisburg-Lebanon-Carlisle, PA	75	7.04	20.5
Hartford, CT	58	2.21	19
Hickory-Morganton-Lenoir, NC	80	16.91	8
Houston-Galveston-Brazoria, TX	68	25.15	27
Huntsville, AL	100	16.83	61
Indianapolis, IN	50	16.44	18.5
Jacksonville, FL	100	21.37	60
Johnson City-Kingsport-Bristol, TN-VA	57	10.1	22

Table 2.9(continued)

CMSA name	Commercialization (%)	Population growth (1990-2000)	Inventor mobility (%)
Kansas City, MO-KS	50	12.2	25.5
Lancaster, PA	80	11.31	14.5
Lexington, KY	100	18.05	24
Los Angeles-Riverside-Orange County, CA	62	12.68	29.5
Madison, WI MSA	33	16.19	14
Melbourne-Titusville-Palm Bay, FL	40	19.36	13
Memphis, TN--AR--MS	50	12.74	16
Miami--Fort Lauderdale, FL	50	21.42	17
Milwaukee-Racine, WI	76	5.13	22
Minneapolis--St. Paul, MN-WI	47	16.94	19
New London-Norwich, CT-RI	0	0.97	20
New York-Northern New Jersey, NY-NJ-CT-PA	49	8.44	25
Norfolk-Virginia Beach-Newport News, VA-NC	67	8.75	25.5
Orlando, FL	50	34.27	25
Parkersburg-Marietta, WV-OH	50	1.39	28.5
Peoria-Pekin, IL	40	2.42	26
Philadelphia-Wilmington-Atlantic City, PA-NJ-DE-MD	50	5.01	26
Phoenix--Mesa, AZ	68	45.27	28
Pittsburgh, PA	53	-1.51	21
Portland--Salem, OR--WA	55	26.3	15
Providence-Fall River-Warwick, RI-MA	100	4.78	20
Provo--Orem, UT	67	39.81	12.5
Raleigh--Durham-Chapel Hill, NC	62	38.85	30
Reading, PA	100	11.03	7
Richmond--Petersburg, VA	0	15.12	13.5
Rochester, NY	46	3.36	10
Sacramento--Yolo, CA	75	21.32	27.5
Saginaw--Bay City-Midland, MI	43	0.94	24.5
Salt Lake City-Ogden, UT	75	24.41	36.5
San Antonio, TX	25	20.2	22
San Diego, CA	63	12.64	30
San Francisco-Oakland-San Jose, CA	52	12.57	36

Table 2.9(continued)

CMSA name	Commercialization (%)	Population growth (1990-2000)	Inventor mobility (%)
Seattle--Tacoma--Bremerton, WA	45	19.68	22.5
South Bend, IN	67	7.49	7.5
St. Louis, MO--IL	50	4.46	13.5
Tampa-St. Petersburg-Clearwater, FL	25	15.86	47.5
Toledo, OH	57	0.66	25.5
Tucson, AZ	67	26.52	16
Washington-Baltimore, DC-MD-VA--WV	58	13.1	25.5
West Palm Beach-Boca Raton, FL	20	31	16
York, PA	80	12.42	34.5
All	55	13.06	22.5

2.5 Analytical strategies

Built upon the data sources I described in the previous section, the major purpose in this study is to analyze the relationship between regional ecology and innovation performance (measured by patenting per capita and commercialization rate) for regions and firms. In other words, I ask, given a patent, does the local ecology predict successful commercialization of the invention at both regional and firm level? If so, then which type of ecology seems most efficacious for which types of firms? The analytical strategy should allow us to test the competing hypotheses examining factors that are likely to contribute to positive or negative externalities in innovation clusters. Here is the outline of my analytical strategies for Chapter 3 and 4.

Chapter 3 examines the impacts of the SME dominated ecology on innovation performance at both regional and firm level. It begins with a baseline of OLS regression model, controlling for the MSA cluster effects. The baseline model estimates patents counts of the metropolitan areas as a function of population, the diversity index (the inverse measure of the standardized Herfindahl index) and the number of patenting firms (with logarithm) in a MSA. The analysis will then add the regional ecology measures (the ratio of small firm patent in a MSA-technology) to see its impact on the patenting activity. Finally, I will add measures of spillovers (e.g., regional mobility, and university R&D expenditures in a region) to see if they predict regional inventing activity and whether they mediate the effect of the regional ecology measure. The same specifications will apply again for the second dependent variable, the regional commercialization rate.

Chapter 3 also examines the effect of regional ecology at the project level. To examine the influence of regional factors on individual firm's innovation outputs, we first realize that individual R&D projects are nested in firms (level 1), and firms are nested in regions (level 2). This provides the intuition that we need to investigate our research questions using a multilevel method. This study will analyze research hypotheses by hierarchical linear models[7] (HLM) to see how regional measures affect the regional and firm level innovative performance.

In this project, since the dependent variable, commercialization, is a dichotomous variable, we did not use current commercial packages such as HLM6, which are not designed to deal with multilevel logit models (Guo and Zhao, 2000). We obtain estimates form the SAS 9.2 Glimmix procedure that can conduct multilevel regression models and it accommodates logistic regression.

Subsequently, Chapter 4 investigates the effects of regional ecology by firm size. Hence, the specifications begin with a baseline HLM regression model estimating the commercialization propensity of triadic patents as a function of project-level characteristics (e.g., patent value, scale of the project, technology fields, age of the patent) and regional level variable (i.e., regional resources and regional ecology). Then, I add

[7] The HLM analyses can account for random effects of regional variables to control for unobserved variation of regions, as well as control for the contemporaneous correlation of dynamic changing relevant to the innovation production function.

the interaction of regional ecology and firm size to the model. Alternatively, I also estimate the models separately for small and large firms and compare the coefficients between two groups.

2.6 Limitations of the data

The GT/RIETI survey provides rich data of inventors in the United States, including detailed information on the process of innovation development based on inventors' experiences, rather than from managers' perspectives. However, we understand its limitation as the following.

First, patent is not the only mean for appropriation of intellectual properties. The propensity of patenting varies by industries and firm strategies, such as secrecy, lead-time, other legal approaches, and complementary manufacturing/services (Cohen et al., 2000). Therefore, the interpretation of the findings can better accurately represent the patent-based industries (e.g., the pharmaceutical industry and computer industry), although they also apply to patenting strategies of the non-patenting industries (e.g., the traditional machinery industries) by using the nation-wide sample.

Secondly, our sample is unlikely to grasp the overall quality of the R&D team, which could be an important factor in predicting the success of the invention. However, I can control for the education background of the respondents and the project-level characteristics, such as the project size (i.e., man-month).

Thirdly, using triadic patents means we focus on patents targeting the global markets (applied for EPO and JPO, and granted in the USPTO). One caveat is the possibility of oversampling commercialized inventions and large firms' inventions because additional costs involved for filing and maintaining patents in multiple countries may filter out low-value or less-promising patents. We expect a higher rate of large firm patents in our data. To test this, we compare the number of patents by firm size across previous empirical studies (see Figure 2.1). According to the statistics from the U.S. Patent and Trademark Office (USPTO), small entities (< 500 employees) accounts for 26% of patent applications in 2000 and 23% in 2003. It suggests that our survey of triadic patent inventors with 20% of Small and medium sized firms is not deviating too much from the USPTO data regarding the share of US granted to small entities. However, the pitfall of using patent as a proxy of innovation is that patenting enforcement is simply one way of appropriating invention or new ideas of firms. As Audretsch and Acs's (1991) research indicates, small and medium sized firms only accounts for 43% of innovations. Despite large and small firms apply different strategies in developing new products or new services, the use of patent data covers firms participate in the conduct of R&D, either in the patenting business (e.g., the chemical and communication industries) or the non-patenting industries (e.g., the machinery industry).

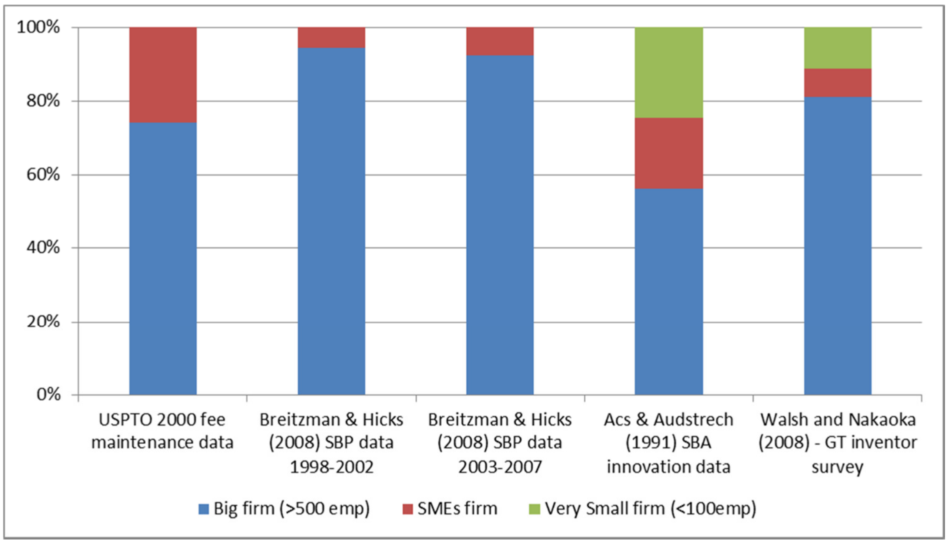

Figure 2.1 Percentage of patents by firm size

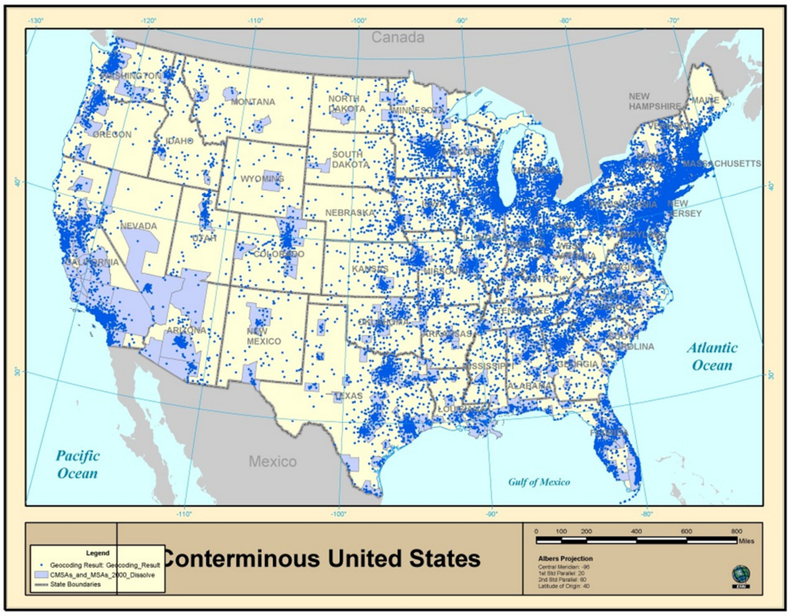

Figure 2.2 A map of patent applications in the US (2000 – 2003)

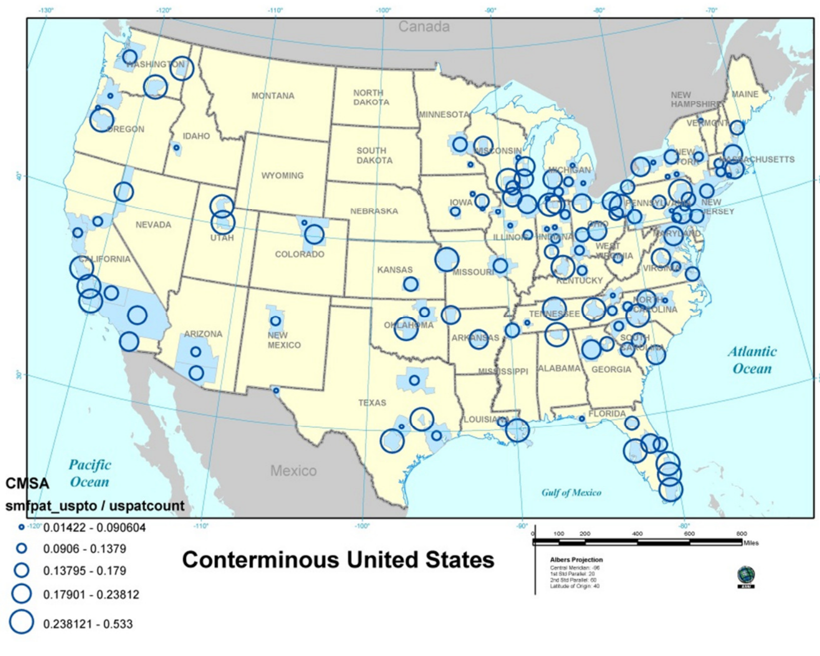

Figure 2.3 Percentage of small firm patents in scales across MSAs in the US

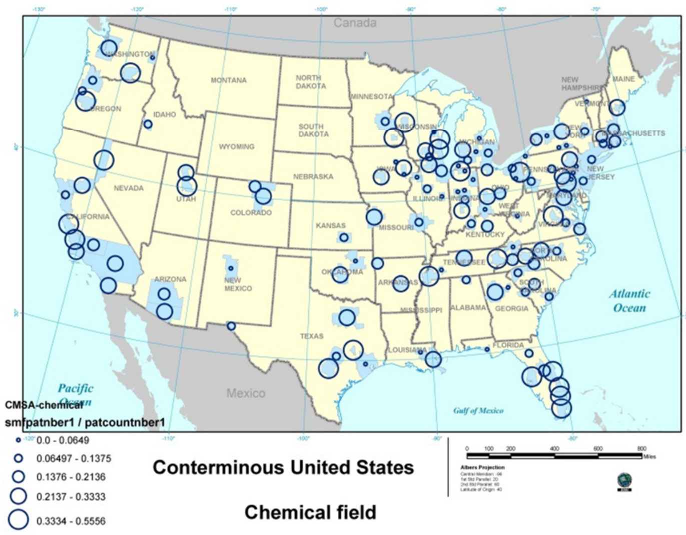

Figure 2.4 Percentage of small firm patents in the chemical field

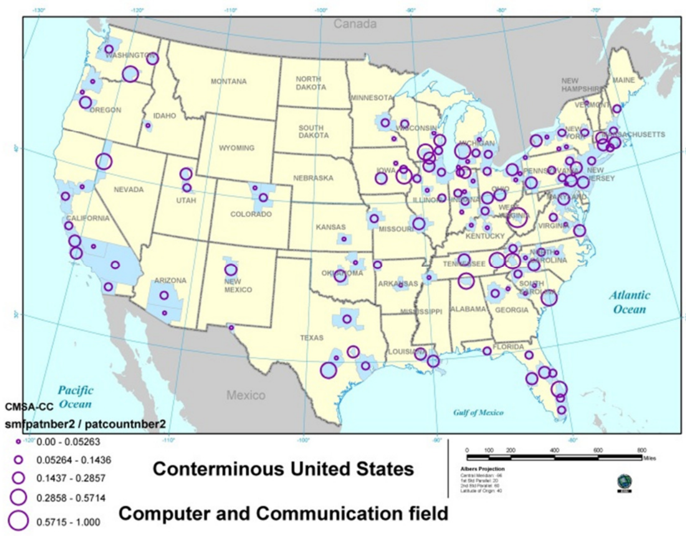

Figure 2.5 Percentage of small firm patents in the computer and communication field

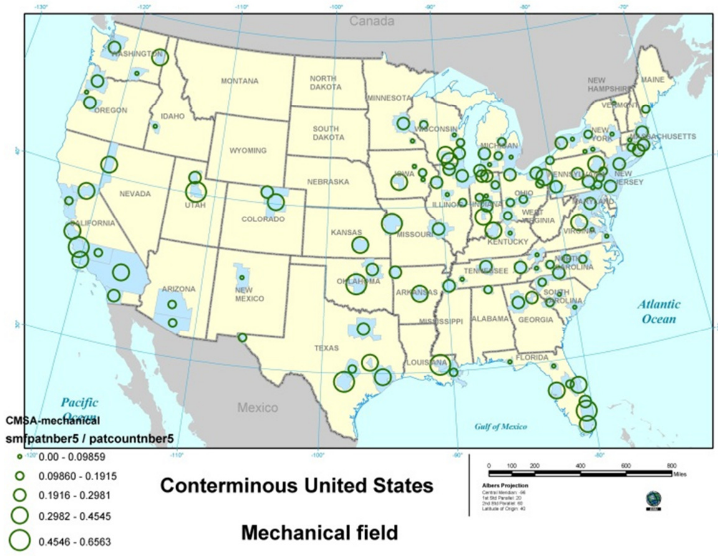

Figure 2.6 Percentage of small firm patents in the mechanical field

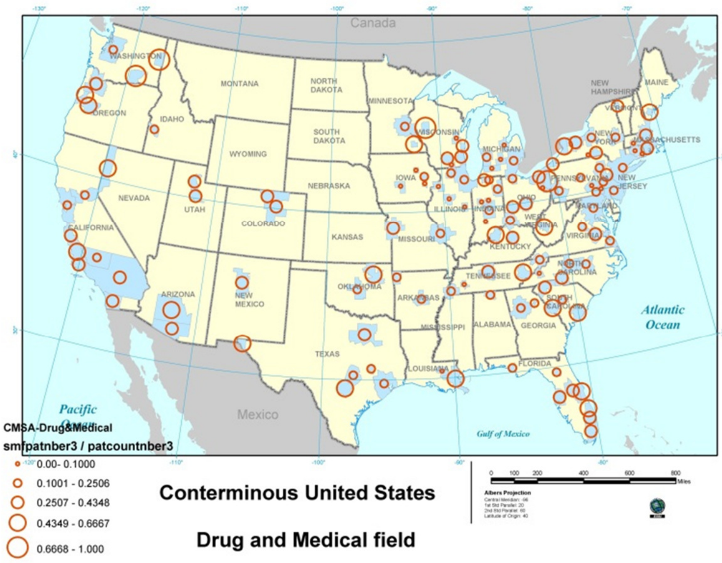

Figure 2.7 Percentage of small firm patents in the drug and medical field

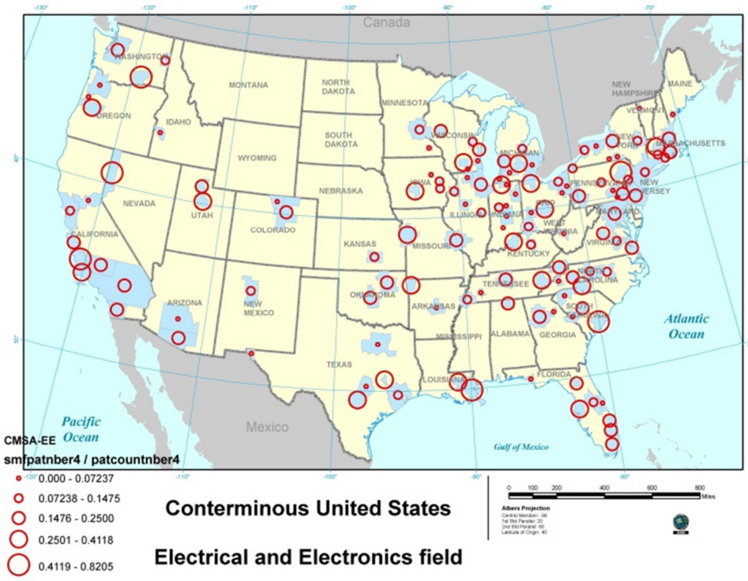

Figure 2.8 Percentage of small firm patents in the electrical and electronic field

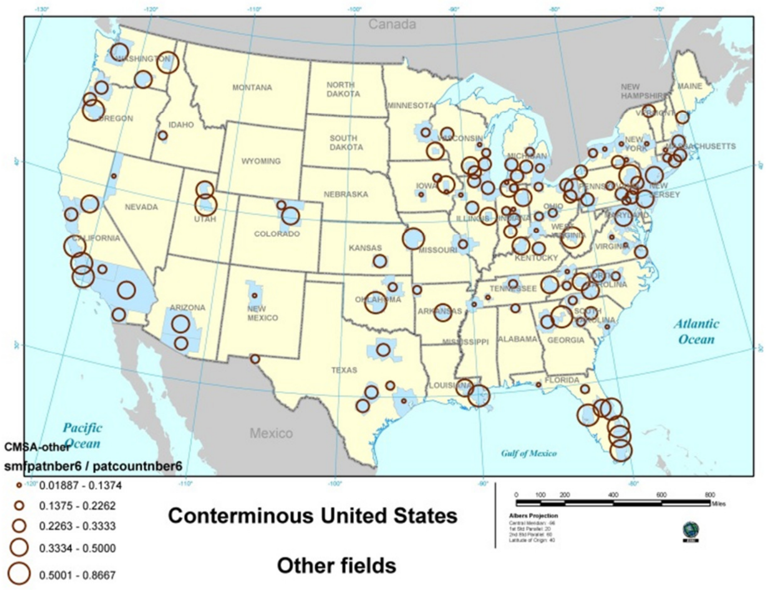

Figure 2.9 Percentage of small firm patents in other fields

CHAPTER 3

REGIONAL INNOVATION PERFORMANCE:

ANALYSES AND RESULTS

3.1 Introduction

This chapter examines two research questions. The first question is the effects of regional resources (based on agglomeration theory) on innovation performance. The second question is to investigate whether the type of concentration (the regional ecology) plays a role in influencing the regional level and firm level innovation performance. Regional ecology is a continuous measure, indicating the proportion of innovative small firms in a region. A higher value represents an ecology dominated by small firms; on the other hand, a lower value represents an ecology dominated by large firms. I operationalize the concept of regional ecology by measuring the share of small firm inventions in a region. In a sense, this measure represents the distribution of firm size in a region, representing the organizational ecology of innovating firms in a region.

The focus of analyses is to investigate whether environmental factors (regional resources and regional ecology) affect the likelihood of innovation performance. Innovation performance is operationalized by two measures, including rates of inventing activity (patent counts at the regional level) and commercialized innovations at the regional level and the project level (we assume that every patent came from a R&D project). In addition, I control for alternative explanations, such as firm characteristics, project characteristics, and technology fields to examine the net effects of regional variables. Following Marshall, this chapter tests the following small-firm dominated ecology hypotheses:

Hypothesis 1a: As the proportion of small firm R&D projects in a region increases, regional innovating activities (patents counts) increase.

Hypothesis 1b: As the proportion of small firm R&D projects in a region increases, regional commercialization rates increase.

The second sets of hypotheses are to test do regional resources (mechanisms of knowledge flows) mediate the relation between regional ecology and regional innovation performance. Hypotheses are:

Hypothesis 2a: The university R&D expenditure in the region mediates the relationship between regional ecology and regional commercialization rates.

Hypothesis 2b: Regional inventor mobility mediates the relationship between regional ecology and regional commercialization rates.

3.2 Regression results

3.2.1 Regional-ecology and rate of inventive activity in the MSA

To predict the regional innovation performance, I first conduct analyses at the regional (MSA) level, regressing regional innovation performance on regional resources and regional ecology.

Table 3.1a represents the OLS regression results predicting rates of patenting activities in a MSA-field. Model 1 is the basic model with variables measuring regional resources. The results show a negative effect of the technology-field diversity on the regional inventive activity rate. Results of Model 1 suggest that the agglomeration of specialized innovative firms is positively associated with rates of patenting activities. The Marshall-Arrow-Romer model (Glaeser et al., 1992) that pertains to external effects of specialized industries in an agglomerated economy is consistent with my findings. The diversification argument of Jacob (Boschma & Frenken, 2009) may provide an alternative model for knowledge searching; my findings reject this standpoint when predicting regional patenting activities.

When adding the regional ecology variable to the model, I found a negative relationship between the proportion of small firm patents in a region and regional patenting activities. This finding contradicts with my Hypothesis 1a that the increasing rate of small firm patents in a region promotes the overall regional inventing activities. I find this result tricky, thus look into what was going on between these variables. The generic problem of the equations of Table 3.1a is the use of ratio variables[8], which is likely to produce bias results and negative correlation between the dependent variable and the ratio variable. Therefore, I take a different approach by not using the ratio variable, but predict the patenting activity as the function of the quintile dummies of the counts of small firm patents in a MSA-field and population in a MSA. Table 3.1b represents the results.

Again, as Table 3.1b Model 1 shows, the regional diversity of technological fields decreases the regional patenting activity. Large MSAs produce more patents. Model 2 adds the number of small firm patents in a MSA-field (with logarithm term) holding all other regional variable constant, and find that the amount of regional patenting activities

---

[8] The equation of inventive activity rate for each MSA-technology would be: $\ln(\frac{P_{all}}{population\ per\ capita}) = \alpha + \beta \ln(\frac{P_{small\ firms}}{P_{all}}) + \gamma Z + error$, where P_{all} = number of patents produced in a MSA-technology dyad, $P_{small\ firms}$= number of small firm patents produced in a MSA-technology dyad, Z = a vector of other variables affecting the rates of inventive activity, and error = statistical residuals. However, as the equation shows, the use of ratio variable as the major independent variable could cause serious problems. Borjas's paper (1980) raised similar issue when predicting wage rate on weekly hour of work. He provided a simple method to avoid the division bias because of the worries of the use of ratio variable in producing negative spurious results.

goes up as there are more small firm patents in a region. To check if this result is correct, I examine a new specification by using the number of large firm patents in a MSA-field in the logarithm term (unreported). I also find positive relationship between large firms patent counts and the regional innovativeness (the coefficient = 0.84 and standard error = 0.14). This result seems to be less useful for answering my first research question. More small firm patents in a region increase the overall regional patenting activity, but so do more large firm patents in a region. As Model 3 notes, the magnitude of coefficient for large firm patents is larger than the coefficient for small firm patents. Alternatively, Model 4 then includes the quintile dummies of counts of small firm patents in MSAs (using the fifth quintile as the reference group). The results show that the fifth quintile group has the highest regional patenting activity compared with the rest of groups. When using the quintile dummies of counts of large firm patents (unreported), results show that the top quintile for large firms has the biggest effect. Based on these results, I have to reject the hypothesis that the concentration of small firm patents stimulates the regional innovative activities, measured by counts of patents. In fact, based on the ratio variable regression and the effects of counts of large firm patents compared to the models with small firm patent counts, the results suggest that the effect of more large firm patents in a region is larger than the effect of small firm patents in a region.

Then, in Model 5, I include two knowledge mechanism variables to the model, one is the amount of university R&D expenditure in MSAs, and the other is the inventor mobility rate in MSAs. The amount of university R&D funding has a positive coefficient on regional patenting activities, but the result is not significant. However, the regional inventor mobility rates are negative to regional patenting activities, which is different from the expectation that circulations of skilled engineers increase the localization of knowledge flow. As prior studies argued, the attraction of skilled labors is positive to the regional innovation (Acs et al., 2002). Regional inventor mobility rates are positively associated with the regional patenting activity if not controlling for other regional variables. Once I add the number of small firm patents to the model, I obtain negative coefficient of the mobility variable. I also find that number of small firm patents is positively correlated with the regional mobility rates, suggesting that a bigger pool of small firm innovation is likely to increase the labor circulation in the local, but the net effect of regional mobility rates is still ambiguous. The conjecture is that the value of labor mobility matters depends on whether they moved within the same field or across fields. In a region of concentrated industry, the effect of labor mobility within the same field could add less or negative value to the R&D project, as well as the creation of new idea. The net value of labor mobility could be less if the movers were constrained by the non-compete agreements when they moved from large firms to small firms. On the other hand, the net value of labor mobility could be positive when the designation company is a large firm. Unfortunately, I do not have data indicating the size of the previous firm of the mobile inventor.

I run Model 5 again by replacing counts of small firm patents with counts of large firm patents in a MSA-field. Results show that the bigger pool of large firm patents also increases innovative activity in a region. Holding the number of large firm patents as constant, I find that regional inventor mobility is positive to regional patenting activities. The correlation coefficient between large firm patents and small firm patents at MSA-field level is around 0.65. It means that both measures capture the innovativeness of the region. More likely, when a region concentrates with large firms, the number of small firm inventions could surge through spinning out from large companies or fragmentizing their technological niches. Acs and Adretsch (1998) had similar argument. They find that in an industry dominated by large firms the level of innovation grows. However, many of these innovative activities occurred in the smaller firms because they can only be competitive if they own innovation.

To conclude, we have to reject the hypothesis that SME dominated regions increase innovative activities at the regional level. The findings do not show strong effects of a bigger pool of small firm innovation determining the overall increase of regional innovation performance. In contrast, a series of my findings show that the concentration of large firms has stronger positive effect on regional innovative performance, particularly the patenting activities. For this section, we suggest that the regional patenting performance links closely with the presence of large firms in a region.

Table 3.1a Results of regressions (DV = rates of patenting activity)

	Model 1 Agglomeration effects	Model 2 Add regional ecology	Model 3
Percentage of small firm patents		-0.011*** (.002)	-0.039*** (.008)
(Percentage of small firm patents)^2			0.0005*** (.0001)
Log(amount of university R&D expenditure) per MSA			
Regional mobility rates			
Diversity of technology field per MSA	-2.568*** (.641)	-1.737*** (.495)	-0.898* (.404)
Log(the number of assignees) per MSA	0.039* (.016)	0.030+ (0.017)	0.052*** (.014)
Intercept	2.464*** (.600)	1.939*** (.473)	1.897*** (.406)
MSA fixed effect	Yes	Yes	Yes
Observations	335	335	335
R-square	0.052	0.131	0.201

+, p<.1; *, p<.05; **, p<.01; ***, p<.001

Table 3.1b Results of regression models (DV = log(counts of patenting activity))

Independent variables	Dependent Variable: log(# of patents in a MSA-field)						
	Model 1 Agglom effects	Model 2 Add #small firm patents	Model 3 Add #large firm	Model 4 Add #small firm patents quintiles	Model 5 Add know flows vars	Model 6 No CA	Mode7 No CA
Log(number of small firm patents in a MSA-field)		0.535*** (.053)	0.198*** (.013)		0.528*** (.059)	0.521*** (.054)	0.501*** (.052)
Log(number of large firm patents in a MSA-field)			0.802*** (0.013)				
# small firm patents -1st quintile(20%)				-0.849** (.255)			
# small firm patents-2nd quintile(40%)				-0.665** (.246)			
# small firm patents-3th quintile (60%)				-0.862*** (.182)			
# small firm patents-4th quintile (80%)				-0.766*** (.117)			
Log(amount of university R&D expenditure) per MSA					0.003 (.006)		0.004 (.006)
Regional mobility rates per MSA					-0.0163+ (.009)		-0.215* (.009)
Diversity of technology field per MSA	-1.384+ (.853)	-2.094** (.799)	-0.205 (.107)	-2.062* (0.826)	-1.798* (.796)	-1.870* (.788)	-1.506* (.754)
% of startups per MSA	0.621 (.521)	0.083 (.266)	-0.065 (.047)	0.343 (.406)	0.142 (.267)	0.218 (.349)	0.279 (.319)
Log(population of year 2000) per MSA	0.779*** (.067)	0.316*** (.062)	-0.022* (.010)	-0.587*** (.076)	0.312*** (.066)	0.297*** (.066)	0.293*** (.069)
Intercept	-4.843*** (.961)	0.459 (0.893)	1.033*** (.117)	-1.310 (1.2127)	0.582 (.925)	0.590 (.911)	0.792 (.933)
MSA fixed effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	552	543	552	552	515	508	472
# of Cluster (MSA)	125	125	125	125	118	119	112
R-square	0.489	0.624	0.991	0.542	0.624	0.581	0.581

+, p<.10; *, p<.05; **, p<.01; ***, p<.001 ; Specifications in this table exclude cases are not located within any MSA and MSAs with less than 20 patents

3.2.2 Regional-ecology and commercialization rates in MSAs

Next, for the second innovation performance measure, I examine whether the concentration of many innovative small firms enhances or mitigates the likelihood of commercialization in addition to traditional concentration (regional resources) measures. To assess this research question, in Table 3.2, I investigate the relationship between regional ecology (the percentage of small firm patents) and commercialization rates in MSAs.

The first model (Model 1 of Table 3.2) tests is to predict regional commercialization rate a function of inventing activity rates and the regional diversity of technology field (i.e., the diversity index of six technology fields in the MSA). Results of Model 1 indicate the inventing activity rate per MSA-field is decreasing commercialization rates, suggesting that a big pool of technologies in a region is likely to have a diminishing effect on commercialization. The diversity measure of technology fields in a MSA is negative but not significant. However, this result is consistent with models in previous sections that diversity decreases patenting activities in a region.

By adding the regional ecology variable, Model 2 reports a positive relationship between the percentages of small firm patents and commercialization rates in a MSA-field, holding other control variables constant. I compute commercialization rates by calculating the mean of any commercial use of the patent invention (*cashin* = 1) for each MSA-field based on the survey data. The coefficient implies that one percent increase in small firm patents will have a 0.4% increase in commercialization rates in a MSA-field. Concerning of ratio variable issues, when predicting counts of small firm patents (with logarithm) on counts of commercialization (with logarithm) in a MSA-field, I obtain similar results. More small firm patents in a region increase the likelihood of the increase of commercialized invention in a MSA-field. If using counts of large firm patents (with logarithm), I find a negative relationship between the number of large firm patents and commercialization rates in a MSA-field. This gives us some confidence that we can accept our hypothesis that SME dominated ecology is positively associated with the regional commercialization activities. Because the denominators of commercialization rates and small firm patents per MSA-field are different, I do not think the use of ratio variable is a big issue here.

To test mediation hypotheses, Model 3 and 4 adds the regional resources variables to the model. In General, a mediation effect (Z) occurs with the following criteria. First, the independent variable X significantly affects the mediator Z. Second, the independent variable X significantly affects the dependent variable Y in the absence of the mediator. Third, the mediator Z has a significant direct effect on the dependent variable Y. Finally, the effect of X on Y shrinks upon the addition of the mediator Z to the model. Additionally, the effect of the mediator Z can be formally assessed by the following equation (Sobel, 1982, MacKinnon et al., 1995). The z-value = $\frac{\beta_a * \beta_b = c - c'}{\sqrt{(\beta_a^2 * se.b^2) + (\beta_b^2 * se.a^2)}}$,

where $\beta_a$is the regression coefficient for the association between X and Z, β_b is the regression coefficient for the association between Z to Y, and se.a/se.b are correspondent standard errors of β_a and β_b. Also, c is the direct effect of the model predicting Y as a function of X without controlling for Z, and c' is the net effect of X on Y adding Z to the model. Mathematically, the total effect c in the binary situation should be equal to

$\beta_a * \beta_b$ + c'. Hence, $\beta_a * \beta_b = c - c'$. Therefore, in my dissertation, the z-value is calculated as the difference of two models and standardized by the standard deviations of the mediated effect.

The big effect here is that regional ecology is positively associated with the regional commercialization rates, consistent with the Marshallian hypothesis. The inclusion of the university R&D expenditure (Model 4) did not change the result significantly, but the inclusion of labor mobility variable (Model 3) decreases the coefficient of regional ecology from 0.0032 to 0.0028 with the mediation test of p-value = 0.12), although it is not a significant change. For a robustness check, Model 5 removes California cases since many critiques mentioned that California is known as the growing State since 1980s, particularly in the Silicon Valley and the Los Angeles region (Fallick et al., 2011; Oden, 2000; Saxenian, 1996). The findings are consistent with models of the full sample.

Table 3.2 OLS regression results of predicting rates of commercialization

	Dependent variable: Commercialization rate in a MSA-field					
	Model 1 Full sample	Model 2 Full sample	Model 3 Full sample	Model 4 Full sample	% of total effect explained by mediator	Model 5 w/o CA
Percentage of small firm patents per MSA-field		0.0032+ (.0018)	0.0028 (.0018)	0.0031+ (.0017)		0.0042+ (.002)
Log(amount of university R&D expenditure) (at the MSA level)				-0.0016 (.003)	3.1% n.s	0.0008 (.003)
Regional mobility rates at the MSA level			0.0049+ (.0028)		12.5% (p=.16)	0.0064+ (.003)
Patent per capita per MSA-field	-0.08*** (.017)	-0.063*** (.016)	-0.053** (.016)	-0.062*** (.016)		-0.03 (.02)
Diversity of technology field	-0.053 (.407)	-0.257 (.411)	-0.487 (.319)	-0.147 (.480)		-0.632 (.413)
% startup	1.102** (.374)	1.001** (.369)	0.812* (.374)	1.156** (.418)		
Intercept	0.596 (.373)	0.731+ (.370)	0.854** (.280)	0.649 (.415)		0.774* (.313)
MSA cluster effect	Yes	Yes	Yes	Yes		Yes
Observations	163	163	152	163		135
# MSA cluster	59	59	56	59		53
R-square	0.078	0.094	0.105	0.096		0.125

+, p<.10; *, p<.05; **, p<.01; ***, p<.001

Models not control for cluster fixed effect produce similar results

Table 3.3 Multivariate regression with percentage of small firm patents as an independent variable and (a) inventor mobility rate and (b) university R&D expenditure as dependent variables, clustered by MSA

	Model 1 (a) labor mobility		Model 2 (b) amount of university R&D expenditure in Year 2002	
	Coefficient	P-value	Coefficient	P-value
Percentage of small firm patents	0.114* (0.053)	2.12		
Percentage of small firm patents			0.103 (.085)	1.20
MSA cluster effect	Yes		Yes	
Observations	152		163	
# MSA cluster	56		59	
R-square	0.061		0.017	

*, p<.05

3.2.3 Regional ecology and commercialization at the patent level

Then, I conduct analyses at the patent level by regressing the commercialization of a patent project on regional variables, firm-level variables, and project-level characteristics. This section emphasizes the role of regional ecology on commercialization while controlling for organizational and project level measures. I adopt the hierarchical linear models (HLM) to investigate associations between variables at different scales. I apply HLM with binary dependent variables, which can be understood in terms of generalized linear modeling approach (McCullagh & Nelder, 1989), but with transformed estimates. According to Raudenbush and Bryk (2002), a basic two level HLM can be written as the following equations:

For level 1: $Y_{ij} = \beta_{0j} + \sum_{p=1}^{p} \beta_{(p)j} X_{(p)ij} + \varepsilon_{ij}$ (1)

For level 2: $\beta_{0j} = \gamma_{00} + \sum_{s=1}^{s} \gamma_{(s)0j} W_{(s)j} + \mu_{0j}$ with $\mu_{0j} \sim N(0, \tau_{00})$ (2)

Equation (1) is the unconditional model with a group-level intercept β_{0j} and is related to a project-level predictor variable $X_{(p)ij}$ by the coefficients of $\beta_{(p)j}$. Because the dependent variable Y is associated not only with the individual *i* observations, but is also associated with the *j* group (Hurlbert, 1984). In HLM, the first level model does not model the overall intercept and slope of the sample, but model around the intercept and slope of each of the level-2 group (j = 1,…,J). Equation (2) depicts the level 2 model, in which γ_{00} is the level 2 coefficient for the intercept, indicating average of log-odds ratio of commercialization across regions. And μ_{0j} is the random effect at level 2, which is assumed to be distributed as multivariate normal with means of zero and variances of τ_{00}. Since we have multiple regional-level variables, we will process a step-by-step inclusion of variables explaining the intercept. At this point, variance in equation (2) is conditional. The inclusion of those contextual variables ($W_{(s)j}$) to equation (2) is to measure the extent of which the average probability of a firm that is commercializing its invention varies among regions due to characteristics (e.g., regional ecology, university R&D expenditure, and regional mobility rates) of the environmental context in which they are located.

In the case of logistic regression, we need a logarithm transformation of predicted probabilities of a binary variableY_j. The combined equation of equation (1) and (2) is presenting in the following form:

$$\ln\left[\frac{p_{ij}}{1-p_{ij}}\right] = \gamma_{00} + \sum_{s=1}^{s} \gamma_{(s)0j} W_{(s)j} + \mu_{0j} + \sum_{p=1}^{p} \beta_{(p)j} X_{(p)ij} + \varepsilon_{ij} \qquad (3)$$

I set both level 1 (project-level) and level 2 (region-level) predictors to be fixed, therefore I do not vary β(p)j. Although I do not test if the coefficient of firm-level variables and regional level variables in my models vary from region to region, HLM allows us to test the between group differences accounted for by having the random intercept, but have the linear model estimators for those fixed variables. This means that we can better obtain precise estimates for the regional effect using multi-level models than the logistic model.

As mentioned in the methodology chapter, results of estimators were calculated from the Proc Glimmix procedure by SAS that can fit a logistic model for multilevel models with random effects. To show that the Glimmix procedure can fit the model with better sensitivity and specificity, I compare the ROC (Receiver Operating Characteristics[9]) between two methods in Figure 3.1. The plot illustrates the ROC curves and areas from both models. The closer the curve follows the left-hand border and then the top border of the ROC space, the more accurate the test is. In addition, the closer the curve comes to the 45-degree diagonal of the ROC space, the less accurate the test is. We can see from the plot clearly that the Glimmix model is a significantly better model than the simple linear effects model fit in PROC LOGISTIC (p <.0001).

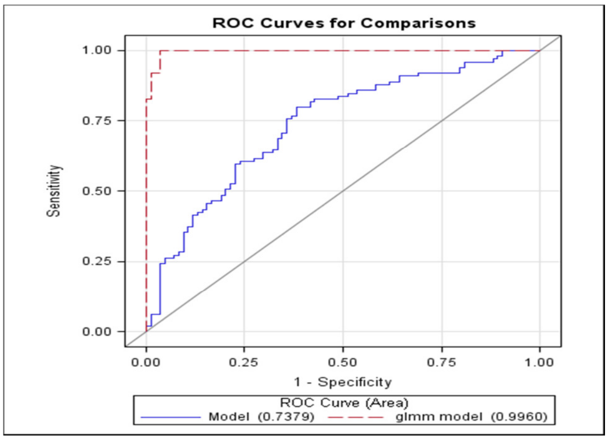

Figure 3.1 Comparison between the Glimmix model and the Logistic model

[9] About the ROC test for the binary dependent variable, see http://support.sas.com/kb/41/364.html

To investigate the role of regional ecology on commercialization at the firm level, I begin with the hierarchical models in Table 3.4. Model 1 is the unconditional model, measuring the variation of between-group means on commercialization. We can interpret the result by first calculating the ratio of commercialized to non-commercialized as $\exp(0.210) = 1.234$. I also run a standard logistic model and the sample ratio is estimated to be $\exp(0.185) = 1.203$, which is the same as the basic sample ratio of 823 commercialized patents to 684 non-commercialized patents. We notice that the standard logistic model is underestimating the ratio by 3% since it does not take into account the clustering effect within groups. Although the difference is quite small, the random effect of intercept (estimate = .102, $p<.05$) suggests there are some additional effects other than the fixed intercept of 0.205, which is average of all block groups (i.e., MSA-field).

Model 2 adds only the firm level and project level variables. The purpose is to test the null hypothesis of no difference between the intercept coefficients across US metropolitan statistical areas (MSAs). However, I cannot reject the null hypothesis (see the bottom row of Model 2). This means that regional variables (level-2 variables) can be treated as fixed measures among US MSAs. Accordingly, in this study, I am not testing whether the regional difference is affected by the random effect of regional ecology since that is not my research focus. However, I think the use of the random effect of MSA-field is necessary to control the possible region-specific effect that correlates with other covariates of the models.

Back to Model 2, as expected, patents owned by large firms are $100\% - \exp(-0.658)*100 = 42.9\%$ less likely to be commercialized. Similarly, projects with higher percentage of basic research components are less likely to be commercialized. Collaborative inventions are 96% more likely to be commercialized than a patent project of in-house R&D, as well as a high ranking (top 10%) patent increases the commercialization of that technology by 220%.

Model 3 adds the regional ecology measure (the percentage of small firm patents) to the regression model and we find that as small firm patent increase 1% in the MSA-field, the commercialization likelihood will increase by exp(0.0159)*100 – 100% = 1.6%. The result is significant at the 10% level. Then, to test whether different sources of knowledge flows mediate the effect of regional ecology, Model 4 includes the university knowledge, measured as the university R&D expenditure in 2002 at the MSA level, to the model. The coefficient of university R&D expenditure is negative, but the result is not statistically significant. In Model 5, I control for regional inventor mobility rates and find that it increases the likelihood of commercialization. One percent increases in the regional mobility rate results in that, the patent is five times more likely to be commercialized. After adding the measure of regional mobility rates, the coefficient of regional ecology drops from 0.0159 (Model 3) to 0.0107 (Model 5), and the effect of regional ecology becomes not significant. The Sobel-Goodman mediation test indicates the mediation effect of regional labor mobility with a p-value of 0.13.

When seeing the positive effect of mobility on commercialization propensity, I wonder whether the effect of inventor mobility is mediated by the strength of non-compete clauses in the States (Fallick et al., 2011; Garmaise, 2009; Gilson, 1998; Marx et al., 2009; Singh and Marx, 2011; Stuart and Sorenson, 2003). In Table 3.4, Model 6 includes a measure of the strength of non-compete enforceability to the model. This measure is created by Garmaise, who surveyed the changes of non-compete laws across states during 1992 and 2005 (Garmaise, 2009). I borrow Garmaise's measure of non-compete clauses ranging from zero to nine. The results of Model 6 show that the net effect of inventor mobility is still significant after controlling for the state-level non-compete enforcement index. In addition, the net effect of non-compete enforcement is negative but not significant. This result suggests that the effect of mobility is not explained by the absence of non-compete clauses (the mediation test is not significant).

However, a second concern is that whether the effects of mobility are greater in a region with low non-compete enforcement, suggesting that the positive impact of the circulation of skill laborers decreases as the strength of non-compete clause increases. This hypothesis is tied closely to Gilson's (1998) argument that disregarding of the non-compete agreement by the California legal system is the reason that explains the high growth in Silicon Valley, rather than by the unique "culture" of the region (Saxenian, 1996). To test whether the mobility of skilled engineers has greater impact in a low enforcement state, I examine the interaction effect of non-compete and labor mobility. In Model 8, I create eight dummy variables from the non-compete enforcement (NCE) measure. NCE0 means no enforcement at all. NCE9 represents the highest level of the enforcement. NCE1 and NCE8 is missing because I have no respondents were in the states with non-compete enforcement scored as 1 or 8. The results show that effect of an increase in the regional labor mobility rates has an overall benefit to States with both high and low enforcement compared with lower mobility rates in regions (see Model 8).

I obtain consistent findings when excluding California cases. Hence, our findings suggest that mobile inventors are performing better not only in a low enforcing state, but also in a high enforcing state, except the highest enforcement level. The results suggest that the effects of skilled mobile inventors are not restricted completely by the non-compete restriction at the state level. Our finding is interesting because the result is very different from previous findings that generally argued that the effect of mobility was mediated by the strength of non-compete enforcement. My study finds that if taking into account the role of regional ecology in shaping the regional mobility, the non-compete restriction does not decrease the value of mobile inventors. This result suggests a scenario that the state level non-compete enforcement does not prohibit labor mobility of skilled engineers because they may leverage the payoff if they change to a new company with the payoff if they stay in the presence of the non-compete regulation. What my results suggest is that the payoff to move and work on something more innovative is bigger than another option. A bigger pool of innovative small firms in a region facilitates this scenario.

Table 3.4 HLM logit results on predicting the commercialization propensity – full sample

	Dependent Variable: Commercial use of the patented invention (Y/N)					
	Model 1	Model 2	Model 3	Model 4	Model 5	
	Level 0	Only firm variables	Add regional variables	Add UNIV effect	Add mobility effect	mediation test
Regional variables						
% SME patents (per MSA-feild)			0.0159+ (.0087)	0.0157+ (.0087)	0.0107 (.0091)	
Technology diversity per MSA			-0.161 (1.365)	0.052 (1.462)	0.182 (1.388)	
Log(Univ RD expenditure) per MSA				-0.0064 (.015)		1.2%n.s.
Inventor mobility rate per MSA					1.755+ (0.941)	33% (P=.13)
Firm variables						
Large firm (>= 500 emp)		-0.658*** (.198)	-0.607** (.200)	-0.612** (.233)	-0.576** (.202)	
Collaboration		0.672*** (.161)	0.669*** (.161)	0.668*** (.161)	0.662*** (.162)	
Co-assignees		0.290 (.510)	.283 (.508)	0.289 (.509)	0.245 (.510)	
# inventors		0.031 (.036)	.035 (.037)	0.035 (.037)	0.043 (.037)	
Inventor-months		-0.0006 (.036)	-0.0007 (.003)	-0.0007 (.003)	-0.0004 (.003)	
High value patents		1.163*** (.217)	1.163*** (.216)	1164*** (.217)	1.148*** (.217)	
% basic research		-.0144*** (.004)	-0.0143*** (.004)	-0.0141*** (.004)	-0.014** (.004)	
VC funding		0.0017 (.005)	0.0018 (.005)	0.0018 (.0047)	0.0016 (.0046)	
Is the patent in the dominant field of the MSA (y/n)		0.117 (.161)	0.219 (.171)	0.209 (.172)	0.183 (.174)	
Intercept – MSANBER mean cashin, r00	0.210*** (.0607)	1.802*** (.500)	1.517 (1.136)	1.443 (1.155)	0.926 (1.183)	
Patent issued year	Fixed	Fixed	Fixed	Fixed	Fixed	
Technology class	Fixed	Fixed	Fixed	Fixed	Fixed	
Subject (MSA)	79	78	78	78	78	
Observations	1507	1056	1056	1056	1056	
Random effect MSA-field mean (std.err)	0.201** (.064)	0.024 (.039)	0.0053 (.035)	0.0067 (.036)	0.0067 (.0307)	

+ p<.10, *P <.05, ** p<.01, *** p<.001

Table 3.4-continued HLM logit model predicting the commercialization propensity of the patented invention – full sample

	Model 6	Mediation test	Model 7	Model 8
	Add NC effect		NC*mobility	NC*mobility
% SME patents (per MSA-field)	0.0067 (.010)		0.007 (.009)	0.0106 (.009)
Technology diversity per MSA	0.643 (1.075)		0.333 (1.557)	0.278 (1.453)
Inventor mobility rate per MSA	1.717+ (1.076)		-0.535 (1.918)	
Strength of non-compete enforcement (NCE) per state	-0.0124 (.0414)	P=0.76 n.s.	-0.185 (.129)	
Inventor mobility*NCE (0-9)			0.605 (.433)	
Inventor mobility* NCE0 (lowest NCE)				1.810* (.928)
Inventor mobility* NCE2				3.219* (1.546)
Inventor mobility* NCE3				3.318* (1.404)
Inventor mobility* NCE4				2.188 (1.403)
Inventor mobility* NCE5				3.005* (1.301)
Inventor mobility* NCE6				3.771** (1.178)
Inventor mobility* NCE7				4.445 (3.237)
Inventor mobility* NCE9 (highest NCE)				0.070 (1.877)
Large firm (>= 500 emp)	-0.526** (.224)		-0.550* (.226)	-0.590** (.203)
Collaboration	0.799*** (.179)		0.802*** (.179)	0.638*** (.163)
Co-assignees	0.716 (.578)		0.747 (.578)	0.286 (.511)
# inventors	0.047 (.041)		0.047 (.041)	0.048 (.037)
Inventor-months	-0034 (.004)		-0.003 (.004)	-0.0005 (.003)
High value patents	1.209*** (.236)		1.187*** (.237)	1.167*** (.218)
% basic research	-0.015*** (.004)		-0.015*** (.004)	-0.015*** (.004)
Venture capital funding	0.001 (.005)		-0.0009 (.005)	0.0013 (.003)
Is the patent in the dominant field of the MSA (y/n)	0.255 (.196)		0.266 (.197)	0.120 (.177)
Intercept – MSA_field mean cashin, r00	0.667 (1.334)		1.571 (1.481)	0.807 (1.218)
Patent issued year	Fixed		Fixed	Fixed
Technology class	Fixed		Fixed	Fixed
Subject (MSA)	78		78	78
Observations	1056		1056	1056

+ p<.10, *P <.05, ** p<.01, *** p<.001

Instrumental variables: population growth rate

To identify the causal effect of regional ecology on firm innovation, we need to address a potential bias coming from the correlation between regional unobservable heterogeneity and regional ecology (the concentration of innovative small firms). To address the endogeneity concern, this study introduces an instrumental variable that is likely to affect the formation of regional ecology, but does not affect the dependent variable and errors. This study chooses population growth as the instrumental variable, suggesting that the population growth of a region increases incomes the pool of employment and size of the market (Glaeser et al, 1995). These niches could be attractive to entrepreneurs; hence, we should expect the increase of population to be positively associated with an increase of local small business (a small firm dominated ecology). Meanwhile, as a second criterion of an instrumental variable, we should expect variation in population growth rates do not correlate with errors in the model correspond to innovation growth. Supposedly, the population growth (1990-2000) in ten years did not fluctuate by the cross-sectional commercialization rates in MSAs (the sample period is during 2002 and 2006).

Table 3.5 presents estimates of the instrumental variable regression. I use Heckman selection correction model for binary dependent variable to address this endogeneity concern. Model 1 shows the results of the probit model, the results are similar with the basic model in Table 3.4. Model 2 reports coefficients of the first step specification, showing that the population growth during 1990 and 2000 in a MSA is positively associated with regional ecology (SME dominance). Because the rate of patenting activity is likely to increase small firm innovation as well, I include it to the first step model. However, patent per capita has no significant impact on the shape of a small firm dominated region. The Wald test is not significant suggesting that the correlation between instrumental variables and error term is not high (chi-square = 0.01, p<.0.94). Model 3 reports the second stage estimates with the regional ecology variable instrumented by the population growth in a MSA. The results show that the Heckman selection model is similar with the simple probit model in Model 1 in which the estimate of the concentration of small firms increases the likelihood of commercialization significantly. Model 3 shows that the second step regression is still robust with the inclusion of population growth. The coefficient of regional ecology still has a significant positive effect on commercialization.

The use of population growth as the instrumental variable may not the optimal choice since intuitively the increase of population should also increase the scale of market, hence we should expect the positive relationship between the population growth rates and the commercialization outcomes. A better instrument, such as the income tax credit, could be a state-level intervention that creates incentives for entrepreneurs and the creations of small firms. Future work should pay more attention to address the endogeneity issue.

Table 3.5 Population growth as the instrumental variable predicting the SME dominated ecology using the Heckman selection correction estimations

(Dependent variable= commercial use of the patented invention, Instrumental variable = population growth)	Model 1	Model 2	Model 3
	Probit Comm	Selection model % Small firm patent	IV probit Comm
Regional-level variables			
%small firm patents	0.010+ (.006)		0.0103+ (.005)
Population growth rate		0.204+ (.125)	
Rates of innovating activity		0.153 (.231)	
Firm level variables			
Large firm (> 500 employees)	-0.361*** (.104)		-0.363** (.104)
Collaboration	0.403*** (.105)		0.403*** (.104)
Co-assignees	0.154 (.272)		0.159 (.273)
# inventors	0.018 (.023)		0.020 (.024)
Inventor-months	-0.0003 (.003)		-0.0002 (.002)
High value patents	0.698*** (.107)		0.692*** (.107)
% basic research	-0.009*** (.002)		-0.009*** (.002)
Any VC funding	0.0013 (.002)		0.001 (.003)
Is the patent in the dominant field of the MSA (y/n)	0.135 (.081)		0.131 (.080)
Rho			-0.108 (1.024)
Wald test	156.2		162.9
MSA cluster adjust	78	78	78
Observations	1056	1058	1058
Chi2 of independent equation			0.01 (p=0.92)

+p<.1, *p <.05, ** p<.01, *** p<.001;

Robustness Check

In addition, I checked several alternative models for robustness checks. As shown in Table 3.7, In Model 1 and Model 2, I remove the California cases and obtain similar results that the effect of regional ecology is positive and significant to commercialization. Regional mobility rate mediates the effect of regional ecology on commercialization in non-California areas as well. The mediation test is significant.

Model 3 shows a model limiting cases in the big metropolitan areas, where the population exceeds 1,000,000 people. Small MSAs are likely to generate biased results due to the small share of patenting activities compared to the national average. Moreover, empirical studies often focus on the prosperous regions with a large population. Hence, I double-check whether the results are consistent in those big MSAs. Again, results are very similar with models shown in Table 3.5 using the full sample.

3.3 Summary

This chapter provides quantitative evidence for researchers to revisit the concept of localized knowledge searching. How do firms search localized knowledge is one of the core themes in the field of agglomeration economy (Breschi and Lissoni, 2001), that the concentration of firms in specialized industries brings a pool of technical knowledge and professions that enhances the use and supply of innovation (Feldman and Florida, 1994). In order to do so, this chapter investigates the innovation production function to test the external effects of "knowledge flow" in certain regional ecologies, controlling for regional resources.

Based on the results, I reach the conclusion that the concentration of firms in traditional definition is positively associated with regional innovation performance, yet was not enough explaining the mechanism of the localized knowledge spillover. Regional ecology on the other hand, represents the institutional structure of types of firms in a region, showing to be a significant factor predicting regional and firm innovation performance. In other words, findings in chapter 3 illustrate the importance of the small firm dominated ecology in enhancing commercialization at the regional level and the firm level. This finding accords with the Marshall-Arrow-Romer model that the concentration of specialized small firms provides positive externalities to the regions. This chapter shows that the knowledge spillovers exist by adding two mechanisms of knowledge flows at the regional level. One is the university R&D and the other is the labor mobility of skilled workers (i.e., patent inventors). The findings are consistent with the prediction that the effects of regional ecology are mediated by the local labor mobility. This mechanism is not a special phenomenon merely occurs in Silicon Valley. Based on the results, I suggest that the ecological structure of a region may also shape the structure of the labor market, which further increases the innovation rate of local firms.

Table 3.6 Robustness check on factors predict commercial use of patented inventions

Dependent variable = commercial use of patented invention	w/o CA	w/o CA	Mediation test	Large MSA ≥ 1 million
	Model 1	Model 2		Model 3
%small firm patents in a MSA-field	0.0220* (.0098)	0.0135 (.010)		0.0171+ (.011)
Diversity of technology in a MSA	-0.631 (1.512)	-0.460 (1.454)		0.248 (1.865)
Inventor mobility in a MSA		3.647* (1.125)	-1.70 (0.003) P = 0.088	
Large firms	-0.490* (.233)	-0.473* (.233)		-0.547** (.209)
Any collaborator (y/n)	0.637*** (.178)	0.629*** (.179)		0.616*** (.174)
Any co-assignees (y/n)	0.152 (.602)	0.090 (.597)		0.118 (.511)
# inventors	0.019 (.039)	0.033 (.040)		0.048 (.040)
# Inventor-months	0.001 (.003)	0.0016 (.003)		0.0012 (.003)
High value (top 10%) patents (y/n)	1.167*** (.243)	1.131*** (.245)		1.170*** (.231)
% basic research	-0.012* (.005)	-0.013* (.0047)		-0.015*** (.004)
Startup firm (y/n)	0.111 (.389)	0.053 (.391)		0.080 (.339)
Is the patent in the dominant field of a MSA (y/n)	0.261 (.182)	0.161 (.184)		0.215 (.195)
Intercept – mean cashin for each MSA-field, r00	1.723 (1.269)	0.994 (1.255)		1.300 (1.497)
Patent issued year	Fixed	Fixed		Fixed
Technology class	Fixed	Fixed		Fixed
Subject (MSA)	73	73		48
Observations	854	854		917
Level 2 intercept	0.027 (0.050)	-		0.023 (.047)

+p<.10; *p <.05, ** p<.01, *** p<.001

Model 3 only includes MSA-field in the big metropolitan areas where the population in year 2000 is greater than 1,000,000 (above 80 percentile of the sample MSAs)

CHAPTER 4

REGIONAL ECOLOGY, FIRM SIZE, AND INNOVATION PERFORMANCE

4.1 Introduction

Chapter 4 addresses the third research question, "How does regional ecology affect innovation performance by firm size?" Deriving from the Marshallian thesis, this chapter investigates whether small firms are better off in the presence of many innovative small firms in a region. The concentration of specialized local firms generates positive external effects that facilitate local networks and production. Previous literature implies that an ecology dominated by small firms creates positive knowledge externalities and cooperative atmosphere, which benefits local firms in particular. Accordingly, this chapter tests Marshallian hypotheses by investigating the effects of regional ecology on commercialization of firms of different sizes. Hypotheses are:

Regional ecology vs. firm size:

> *Hypothesis 3a: As the percentage of small patents increases (toward SME - dominated ecology), firms increase the likelihood to commercialize their patented inventions.*
>
> *Hypothesis 3b: The difference in commercialization propensity between large and small firms is larger in a SME dominated ecology than in a large-firm dominated ecology.*

External knowledge sources vs. firm size:

Hypothesis 4a: The effect of university knowledge on firm's probability to commercialize is moderated by firm size. (Small firms benefit more from the presence of universities with high R&D budgets than large firms do.)

Hypothesis 4b: The effect of inventor mobility on firm's probability to commercialize is moderated by firm size. (Small firms benefit more from high labor mobility than large firms do.)

4.2 Regression results

To test how the effect of regional ecology differs by firm size, I first introduce interaction terms to examine if the increase in the ratio of small firms' innovation makes small firms more innovative in terms of the commercial use of their patented inventions. The use of interaction terms allows us to investigate whether the benefit of the concentration of small firm patents is contingent on the size of firm.

Alternatively, I also run separate models for each group (one group is for large firms and the other group is for SMEs). In this case, I compare all parameter coefficients between large firms and SMEs, particularly the regional ecology variable and knowledge mechanism variables. The purpose of having separate models is to test whether large firms and SMEs are affected differently by the ecological contexts, as well as the use of regional knowledge sources.

4.2.1 Interaction term of regional ecology and firm size

To begin, Table 4.1 shows the results of the HLM regressing the interaction term of regional ecology and firm size on firm's commercialization propensity. The dependent variable is a binary variable, measured by any commercial use of the patented invention. The interaction term is the multiplication of the regional ecology measure, a continuous variable, and the firm size, a dichotomous variable. Model 1 shows that the effect of the interaction term of firm size and the regional ecology index is positive to commercialization, suggesting the presence of many innovative small firms are likely to be beneficial more in large firms. However, the statistics test is not significant. Although the number of small and medium sized firms is disproportionally smaller than large firms are in the sample, I do have enough cases to conduct the test. The weak result makes me wonder if the benefit of being in a SMEs dominated ecology does not have uniform effect over the ratio of small firm inventions on commercialization outcomes. In other words, we should expect to see innovation performance of firms in different sizes disproportionally innovate more or less in different regional ecologies. To investigate this, Model 2 adds interaction terms of firm size and quintile dummies of the percentage of small firm innovation. The results suggest that large firm's propensity to commercialization decrease in ecologies with low percentage of small firm patents. In the highest quintile, large firms are positively associated with commercializing their inventions than their small neighbors are, however I do not find this result significant as Model 2 shown. This finding suggests that small and large firms could both benefit from an ecology concentrated with small firms.

One potential scenario is that the large firm is a technology-buyer and the concentration of small firms facilitates this need. Because of the constant supply of specialized technologies from the region, large firms increase the likelihood of commercializing its own technologies internally along with other local inventions. To test this, I then change the dependent variable from firm's propensity to have any commercialization to the propensity to in-house commercialization (Model 3). The results clearly reveal large firms' advantage of being in a small firm concentrated ecology. It means that large firms are more likely to commercialize their patented inventions purely for internal use in a region with high percentage of innovative small firms.

Due to the non-linear nature of the logit model, many scholars argued that we cannot interpret the interaction coefficients in logit models directly; instead, researchers should investigate the marginal effects of the interaction term (Ai and Norton, 2003; Hoetker, 2007; Long, 2009). However, the drawback of interpreting a marginal effect of the independent variable is somewhat "limiting its utility for answering the questions of substantive interest that motivate statistical analysis in the first place" (Zelner, 2009, p1336). Therefore, to better interpret the results of interaction terms, I apply King et al (2000), and Zelner's (2009) approach to provide a graphic presentation of the interaction effect.

In recent years, scholars in social science have begun to address the interpretive issues of conducting logit, probit, and other nonlinear models by applying the simulation-based approach, developed by King and his colleagues (King et al., 2000; Zelner, 2009). To show the simulated predicted probability for large and small firms, I use the Stata command known as "CLARIFY" that was developed by King and his colleagues to plot the results based on 1,000 simulations[10]. Figure 4.1 illustrates the results. The x-axis is the percentage of small firm patents in a region (0 to 1), and the y-axis is the predicted probability of commercialization (0 to 1). The blue line indicates the predicted probability on commercialization for small and medium sized firms. The red line indicates the predicted probability on commercialization for large firms. Figure 4.2 plots the simulated difference in predicted probability between large and small firms.

In summary, these two figures indicate that at the right end of the graph, large firms have lower probability on commercialization than small and medium sized firms. However, the gap is shrinking as the percentage of small firm patents in a region increases. The line for small firms is pretty level, suggesting types of regional ecology

[10] I draw Figure 4.1 and 4.2 based on the simulation code written by Bennet Zelner (2009).

might not affect small firm's commercialization propensity. As Figure 4.1 shows, when the percentage of small firm patents excesses 45%, large firms are having higher expected commercialization propensity than small firms. However, we should have the caveat in mind that we see the overlap in the 95% confidence intervals of the two groups at the right tail of the graph (Figure 4.1). This raises suspicion that we should be careful making strong conjectures.

Table 4.1 Logistic HLM regression model predicting regional ecology on commercialization (with interaction terms)

	Model 1 DV: any kind of commercialization	Model 2 DV: any kind of commercialization	Model 3 DV=pure in-house commercialization
Regional Ecology % SME patents in a MSA-field	0.0065 (.017)		
Large firm (> 500 employees)	-0.756* (.392)		
Large firm * % SME patents	0.0091 (.018)		
Large firm * % SME patents 1st quintile		-0.533 (.446)	0.273 (.440)
Large firm * % SME patents 2nd quintile		-0.744*** (.225)	0.038 (.119)
Large firm * % SME patents 3rd quintile		-0.560* (.239)	0.317 (.225)
Large firm * % SME patents 4th quintile		-0.493 (.335)	0.556+ (.345)
Large firm * % SME patents 5th quintile		1.029 (1.185)	2.411* (1.191)
% startups in a MSA	1.236 (1.537)	1.553 (1.526)	-0.415 (4.453)
Diversity of technology field in a MSA	-0.099 (1.468)	0.665 (1.461)	1.228 (1.391)
Collaboration	0.605*** (.164)	0.616*** (.165)	0.358* (.159)
Co-assignees	0.071 (.506)	0.083 (.508)	-0.494 (.516)
# inventors	0.049 (.037)	0.048 (.037)	0.067+ (.036)
Inventor-months	-0.0002 (.003)	-0.0002 (.003)	-0.0008 (.003)
High value patents	1.185*** (.217)	1.187*** (.218)	0.472* (.194)
% basic research	-0.0127*** (.004)	-0.0127 (.004)***	-0.0148** (.005)
In the dominant technology (y/n)	0.216 (.174)	0.174 (.168)	0.215 (.167)
Intercept: MSA-Field mean of cashin, r00	1.618 (1.252)	1.217 (1.231)	-0.772 (1.16)
Patent issued year	Fixed	Fixed	Fixed
Technology class	Fixed	Fixed	Fixed
Subject (MSA)	263	78	78
Observations	1066	1066	1027
Level 2 intercept	0.0292 (.047)	0.0164 (.0414)	-

+p<.10, *p <.05, ** p<.01, *** p<.001

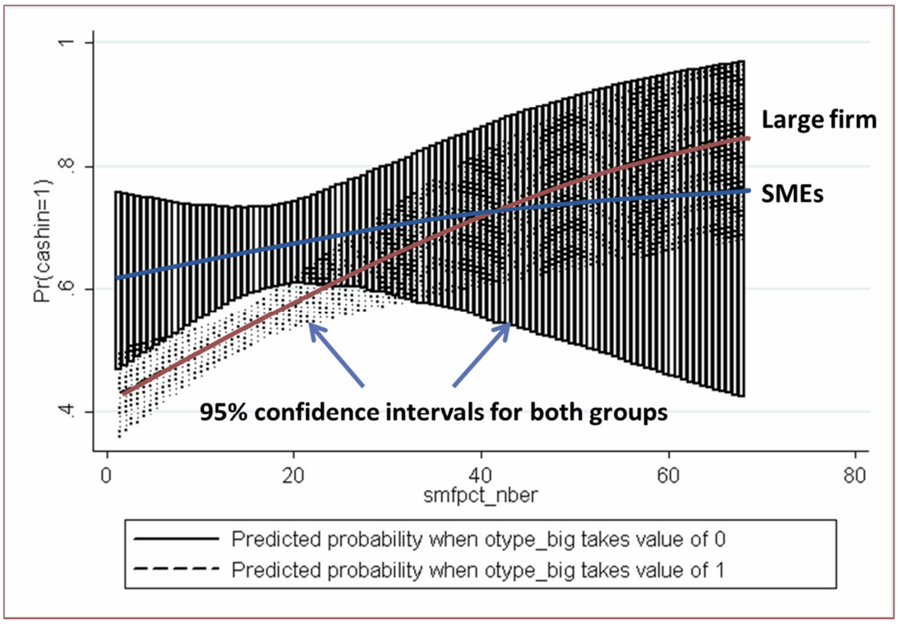

Figure 4.1 Simulated predicted probability of commercialization by firm size

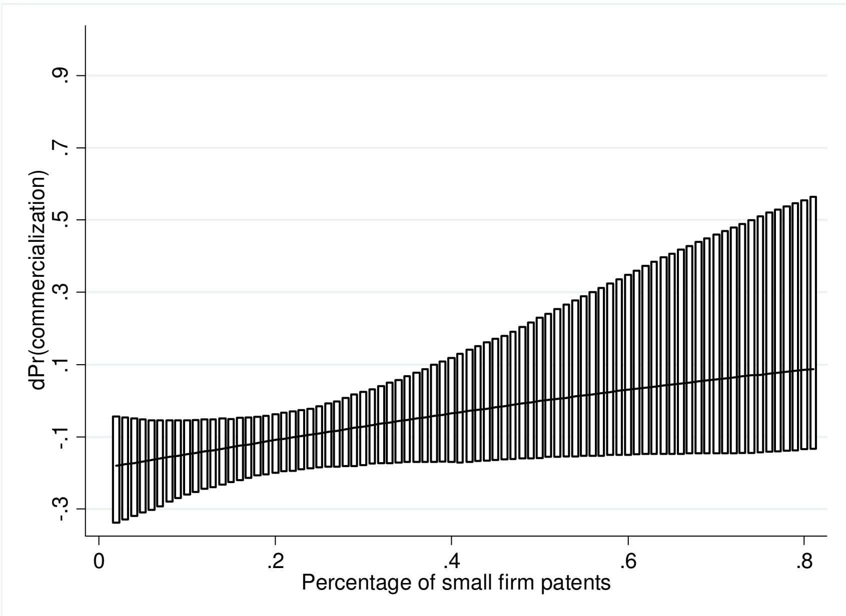

Figure 4.2 Effects of increasing the percentage of small firm patents for ”Large Firms”

Note: The y-axis is the simulated difference in predicted probabilities of commercialization at the current level of %small firm patents (x) with the lower and upper bounds of the 95 percent two-tailed confidence interval

4.2.2 Separate models by firm size

Next, we want to investigate whether the entire models (all parameters) are significantly different from the large firm group and the small firm group. To do so, we separate the analytical model by firm size (large firms and SMEs). Table 4.2 shows the results. Model 1 to Model 3 is predicting the commercialization propensity in large firms. Model 1 shows that the presence of higher concentration of innovative small firms is positive to commercialization in large firms. After controlling for mechanisms of knowledge flows, university knowledge and labor mobility, we find that the effect of regional ecology drops from 0.0161 to 0.008 (from Model 1 to Model 2) and becomes not significant in models of large firms. The results indicate that large firms receive positive benefits from new hires, but negative benefits from the use of university knowledge, albeit not significant. Further, I conduct the mediation test, and the results suggest that labor mobility (critical ratio = 1.546, p = 0.105) is very likely to be the mediator between regional ecology and commercialization for large firms. Again, project level characteristics were used for control variables. For large firms, collaborative patents and high value patent inventions are positively associated to commercialization of the patented invention, which is similar with what we observed in Chapter 3. I obtain similar results while excluding California (Model 3).

Model 4 to Model 6 is predicting commercialization for small to medium sized firms (SMEs). In Model 4, the coefficient of regional ecology variable is negative but not significant at the 5% statistical level. Results in model 5 indicate that the research capacity of universities in the region has a negative impact on small firm's commercialization ability, again is not significant. The university effect has a negative and significant effect on commercialization when we exclude California cases, suggesting that the relationships between universities and small firms are not increasing commercialization as the local practitioners predicted.

To interpret Table 4.2 with careful lens, I compare parameter coefficients across groups in the following paragraphs. As Table 4.3 presents, the last column reports the Wald chi-square statistic for testing the difference between coefficients of large firms and SMEs. According to Allison (1999), the formula for the chi-square statistic is: $\frac{(b_L - b_{SME})^2}{[s.e.(b_L)]^2 + [s.e.(b_{SME})]^2}$, where $b_L$is the coefficient for large firms, b_{SME} is the coefficient for small to medium sized firms and s.e. is the estimated standard error. The findings show that the coefficients of shares of small firm patents in a region are significantly different at the 5% level (chi-square value= 2.90, p<.01).. Based on this result, we can say that each additional 1% increase in small firm patents will increase the odds of commercialization of about 2% for large firms and decrease the odds of commercialization of 6% for SMEs. This 8% difference suggests that large firms get a better payoff from the concentration of small firms than SME. In addition, the effect of having any collaborative partner in the patent project is greater for large firms than it is for small and medium sized firms.

4.3 Summary

In this chapter, I conduct several statistical analyses to illustrate two different innovation production function, one for large firms, and the other for small firms. In general, this study illustrates that in general SMEs have higher commercialization likelihood than large firms. Results of the interaction models show that large firms' propensity to commercialization is increasing as many innovative small firms concentrate the regions. In other words, large firms have higher commercialization pay off in a small firm dominated ecology. The results reject my third hypothesis that small firms are likely to benefit in an ecology dominated with many small firms. I have found the opposite results, which are partially consistent with the power dynamic thesis (Christopherson and Clark, 2007). This suggests that large firms are in better positions to access specialized labor and local resources, but small firms suffer from the power differential in a region. I found only the first part of their argument true in our data. I have also presented this result by providing a graphic presentation of the simulated predicated probability figure (Figure 4.1). We can see the gap of predicted probability of commercialization between large firms and SMEs is decreasing as the percentage of small firm patents increases.

However, the results should be carefully interpreted because of the limitation of a small data cases. Because the sample size of large firms is four times more than the sample size of SMEs, the worry is the error variances are not equal across groups. This assumption is more severe in the logistic regressions than in the OLS models since we are more likely to get biased standard errors and the parameter coefficients (Hoekter, 2007; Allison, 1999). Due to this concern, the regression results (unreported) remain similar if I randomly select 30% of the large firm sample to match with the case of SMEs. However, this study understands that limiting the dataset to a smaller subset does not solve the problem that the residual variances are different across two groups (Williams, 2009). Hence, future study should pay more attentions on addressing this issue.

In conclusion, the findings address the importance of the regional contexts. Regional ecology, particularly the SMEs ecology, shapes the concentration of knowledge resources and knowledge flows in the region. Again, similar to what I found in Chapter 3, regional labor mobility might mediate the relationship between SMEs dominated ecologies and commercialization. Accordingly, the next research question I would like to investigate is whether the effect of labor mobility is within-industry or across-industry.

Interestingly, although my findings confirmed that university knowledge had no direct impact on commercialization, the negative coefficient of university R&D capacity is significantly smaller for large firms than for SMEs. This means that large firms suffered less from the basic knowledge of local universities. In other words, large firms are good at absorbing university knowledge compared to SMEs as previous studies expected (Cohen et al., 2002; Feldman, 1994). What I find is that large firms undertake less negative externality from using local university knowledge than small firm.

In addition, large firms have greater use of collaborations than smaller firms do. This result makes us wonder the mechanism behind it. The first explanation is that large firms collaborate more than small firms (Saito and Okamuro, 2006). However, in our data, inventions of large firms did not collaborate more than that of small firms (30% for large firms and 43% for small firms with any collaboration partners). The second explanation is the lack of internal R&D capability in small firms. Small firms usually collaborate with universities and their customers or suppliers. The university-industry collaboration is likely to be a consulting tasks for small firms since small firms have less capacity to pursuit long-term projects. The worst situation is that the large-n-small firms' collaboration might be hierarchical. Therefore, small firms were in the position of taking assignments, but not able to decide the direction of the project. All these reasons could limit the benefits of the inter-organization collaborations for SMEs.

To conclude this chapter, I have answered the third research question by showing that large and small firms are not receiving the same benefits in the same region, in which regional ecology predicts this difference. However, I do not find significant moderation effects that large firms capture more skilled mobile-laborers than SMEs in the same region. Thus, due to the weak results, I do not have enough evidence to support Hypothesis 4a (university knowledge flow) and Hypothesis 4b (labor mobility knowledge flow) that the firm size moderates the effects of external knowledge sources on commercialization.

Table 4.2 Separate HLM logit models by firm size on commercialization

	Large firms			SMEs		
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
	Full	Full	w/o CA	Full	Full	w/o CA
%SME firm patents in a MSA-field	0.0161+ (.009)	0.008 (.010)	0.013 (.011)	-0.062 (.045)	-0.004 (.027)	-0.022 (.033)
Diversity of technology fields in a MSA	0.635 (1.598)	1.419 (1.64)	0.901 (1.711)	-5.298 (4.399)	-1.365 (4.866)	-7.335 (6.785)
Log(univ RD expenditure) in a MSA		-0.011 (.016)	0.007 (.017)		-0.096 (.063)	-0.110+ (.078)
Mobility rate in a MSA		2.865** (1.029)	4.021** (1.222)		2.202 (3.52)	0.394 (5.252)
Collaboration	0.647*** (.182)	0.640*** (.178)	0.643*** (.198)	-0.053 (.477)	-0.262 (.414)	-0.157 (.510)
Co-assignees	0.125 (.623)	0.084 (.588)	0.278 (.666)	-0.038 (0.958)	-0.186 (.965)	-2.251 (1.762)
# inventors	0.042 (.040)	0.058 (0.040)	0.036 (.043)	-0.025 (.102)	0.023 (0.104)	-0.007 (.122)
Inventor-months	0.003 (.003)	0.002 (.003)	0.004 (.003)	-0.004 (.007)	-0.004 (.008)	-0.001 (.010)
High value patents	1.321*** (.256)	1.228*** (.248)	1.245*** (.276)	0.832+ (.452)	0.853+ (.446)	1.345* (.638)
% basic research spent in a project	-0.014** (.004)	-0.015** (.004)	-0.013* (.005)	-0.0005 (.011)	-0.001 (.011)	0.011 (.018)
In the dominant technology (y/n)	0.091 (.186)	0.119 (.209)	0.064 (.199)	1.329* (.572)	1.243* (.589)	1.013 (0.674)
Intercept – MSANBER mean cashin, r00	85.77 (113.4)	79.63 (113.71)	59.25 (124.3)	282.38 (285.6)	279.86 (289.51)	570.85 (358.59)
Patent issued yr	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed
Technology field	Fixed	Fixed	Fixed	Fixed	Fixed	Fixed
Subject (MSA)	77	77	72	47	47	43
Observations	864	864	728	180	180	130
Random effect						
MSA-Field mean – muoj-large	0.035 (.058)	0.0073 (.041)	0.056 (.100)	0.244 (.412)	0.322 (0.452)	0.870 (0.833)

+p<.10, *p <.05, ** p<.01, *** p<.001

Table 4.3 Chi-square tests of HLM logit regressions by firm size – basic model

	Large		SMEs		Ratio	Chi-square for difference
	Coefficient	SE	Coefficient	SE		
%small firm patents	0.0161+	0.009	-0.062	0.045	-0.26	2.90**
Diversity of technology field	0.635	1.598	-5.298	4.399	-0.12	1.61
Collaboration	0.647***	0.182	-0.053	0.477	-12.21	1.88*
Co-assignees	-0.125	0.623	0.038	0.958	-3.29	0.02
# inventors	0.042	0.040	0.025	0.025	1.68	0.13
Inventor-months	0.003	0.003	-0.004	0.007	-0.75	0.84
High value patents (top 10%)	1.321***	0.256	0.832+	0.452	1.59	0.89
% basic research	-0.014***	0.004	-0.0005	0.011	28.00	1.33
In the dominant field	0.091	0.186	1.329*	0.572	0.07	4.24***
Intercept	85.77	113.4	282.38	285.5	0.30	0.41

+p<.1, *p <.05, ** p<.01, *** p<.001

Table 4.4 Chi-square tests of HLM logit regressions by firm size – labor mobility model

	Large		SMEs		Ratio	Chi-square for difference
	Coefficient	SE	Coefficient	SE		
%small firm patents	0.008	0.010	-0.004	0.027	-1.86	0.18
Diversity of technology field	1.419	1.737	-1.365	4.866	-1.04	0.29
Log(univ RD expenditure)	-0.011	0.017	-0.096	0.063	0.11	1.70+
Mobility rate per MSA	2.865**	1.029	2.202	3.52	1.30	0.03
Collaboration	0.640***	0.178	-0.186	0.528	-3.44	2.20*
Co-assignees	0.084	0.588	0.533	1.338	0.16	0.09
# inventors	0.058	0.040	0.023	0.134	2.52	0.06
Inventor-months	0.002	0.003	-0.004	0.010	-0.50	0.33
High value patents (top 10%)	1.228***	0.248	0.853+	0.553	1.44	0.38
% basic research	-0.015	0.004	-0.001	0.013	15.00	1.06
In the dominant field	0.119	0.209	1.243*	0.830	0.10	1.72+
Intercept	79.63	113.7	279.86	289.5	0.30	0.41

+p<.1, *p <.05, ** p<.01, *** p<.001

CHAPTER 5

CONCLUSION AND POLICY IMPLICATIONS

5.1 Introduction

Overall, this research aims to understand the determinants of innovation performance in US firms by studying organizational and regional contexts. The overarching research inquiry is to understand the mechanism of information flow process to explain regional innovation performance. I propose to pay more attention on the role of regional ecology on regional innovation performance. This study focuses on three research questions. First, what is the role of regional resources on regional and firm's innovation performance? Secondly, does regional ecology, the distribution of types of firms in a region, enhances or reduce the innovation performance? Particularly, do small firm dominated ecologies promote regional innovation performance? Finally, does regional ecology benefit large and small firm equally?

While a vast majority of literature associates agglomeration of firms (geographic proximity) with economic growth, studies on types of firm concentration and regional innovation capacity are relatively scarce. My theoretical contribution is to fill this gap by presenting insight into regional ecology issues, which I mean the mix of firm size in a region. Its impact on the innovation production function for regions and firms.

While reviewing previous literature on regional economics, existing measures, such as the concentration indices, do a good job in capturing the effect of a scale economy and geographic proximity. However, we propose a new measure, regional ecology, to take into account the intra-regional ecological structure. The concept of regional ecology in this study is operationalized as the distribution of large and small firms in a region. For regional policy makers and firm decision makers, in addition to existing regional measures, the regional ecology measure can better explain the mechanisms of the information flow process of collocated firms. Secondly, the empirical contribution is to study innovation performance in terms of both inventions and the commercialization of innovations (i.e., new product and new process), more than counting patents (i.e., new ideas and concepts), which is a relatively new approach. Ihe above research questions will be investigated using a quantitative methodology by focusing on high-tech firms engaged in transforming patent inventions into commercialized innovations.

Data used in this study is a survey of 1,919 triadic patent inventors, which is enriched by detailed information of the use of the patent invention and the characteristics of the particular R&D project. As the major explanatory factor, I collected the regional ecology measures based on a dataset of USPTO patents filed from 2000 to 2003. The empirical evidence demonstrating that regional ecologies are not only spatial containers of firms but act as different kind of social structures influencing knowledge flows and innovation performance for region and firms. More interestingly, organizations of different sizes are shown to react differently under the influence of differing regional contexts.

The following sections put together the findings and comments to formulate the broad implications of these results. The findings will be based on the empirical results and will then discuss the theoretical interpretation and policy implications. The final section of this chapter addresses the limitations of the research and suggests possible topics for future research, as well as policy implications.

5.2 Summary of findings

To answer the first research question, I investigate the relationship between regional resources and regional innovation performance, measured using regional innovating activity and regional commercialization data. In operational terms, previous studies suggest that key regional resources can be defined as classic institutional capacities, including clusters of firms in related industries, industrial R&D agglomerations, and university R&D (Feldman, 1994; Feldman and Audretsch, 1999a). This study defines regional resources as the agglomeration of innovative firms and the diversity of firms in different technology fields. As part of the regional resources, university R&D expenditure is one of the knowledge flows mechanisms, as well as the regional inventor mobility rates. In Chapter 3, the results show that the agglomeration of innovative firms and the cluster of specialized firms are positively associated with regional patenting activities. The current findings support Glaeser's argument that follows the Marshall-Arrow-Romer model (Glaeser et al., 1992) regarding to the external effects of specialized industries, but not necessarily a local monopoly model. This finding is contrary to some empirical results that suggest that diversification of industries facilitates the exchange of complementary knowledge across firms in different fields (Boschma & Frenken, 2009; Feldman & Audretsch, 1999b).

To answer the second research question, we investigate the net effect of regional ecology on innovation performance in addition to those traditional regional resources variables. The results show (see Table 3.1) a positive relationship between the number of small firm patents in a region and regional patenting activities, controlling for the population in each metropolitan area. However, the effect of large firm patents is also positive and strong in predicting the regional patenting activities. Unfortunately, we cannot accept my Hypothesis 1a that the increase of small firm patents in a region promotes the regional inventing activities, or the positive effect of a small firm dominated ecology on commercialization. In contrast, the findings suggest the Anchor-tenant model is likely to increase the level of innovation at a regional level. We found that the magnitude of coefficient for large firms is bigger than that of small firms. To discuss the benefits of the concentration of large firms in a region, Agrawal and Cockburn (2003) have shown that anchor-tenants have the effect of knowledge spillovers when collaborating with local universities. Similarly, Feldman (2003) argues that large anchors are the center of innovative ideas. The presence of large firms enhances regional innovation systems because the increase of spinning off new local innovative firms and the attraction of competitive labors (Niosi and Zhegu, 2005; Link et al., 2003). This result remains the same even if we control for counts of large firms in a MSA aggregated from the GT/RIETI survey (we still get a positive effect). This finding is different from the study by Acs and his colleagues (2002). They found that the presence of large firms is negatively associated with the regional innovation and patenting activities by using firm data in 1982. My explanation is that firms' competitive strategies on the conduct of innovation have changed a lot in the past 20 years, from an internal-focus R&D approach to a more external-knowledge seeking approach as Chesbourgh argued. Hence, my study is a revisit to this research question whether the concentration of large firms may or may not create the regional innovation system. More importantly, the findings suggest that the small firm dominated regions may not predict the increase of patenting activities. I conclude that in a region is dominated by large firm innovations, the level of patenting activities increases.

For the second innovation measure, the findings suggest that the SME dominated ecology is positively associated with regional commercialization rates, which is consistent with the Marshallian hypothesis (support Hypothesis 1b) about the positive externality of the cluster of specialized small firms. Using a robustness check, the effect when controlling for large firm patents on commercialization at the regional level is negative. These results do not change significantly when we use the 250 employees as the cutoff point when creating the firm size dummy. The results remain the same when excluding California cases. The results are also similar when we use population growth as the instrumental variable to predict the formation of an ecology dominated by small firms.

It is important to respond to the mixed results when predicting innovation performance using two different innovation measures. It can be suggested that the patenting and commercialization mechanisms are not the same. Although firms patent their inventions to capture the rent of commercial monopoly, it is not necessary true that firms can successfully commercialize their patented inventions. Many patenting activities were used for strategic reasons, such as protecting existing products and technologies, and easing R&D collaborations. This is confirmed by Agrawal, Cockburn, and Rosell (2010), where inventors employed by the large firms in a company town cited their own patents repeatedly and developed patents based on their prior inventions. On the other hand, small firms are less myopic in patent citations, which is an important characteristic of small firm patents. My conjecture is that the external citations of small firm patents increase commercial value. Small firms' patents recombine the diverse ideas elsewhere, thereby making their inventions more marketable than large firms' patented inventions. At the regional level, the concentration of small firms increases the circulation of marketable ideas, as well as the chance to commercialize those recombined technologies.

In addition, this study found evidence that the regional mobility rates of inventors are likely to mediate the effect of regional ecology when using mediation tests. This finding might provide some evidence supporting Hypothesis 2, where the concentration of innovative small firms in region provides a positive environment for labor flows. Consistent with prior studies, the positive labor mobility loop in Silicon Valley promotes the knowledge spillover process (Saxenian, 1996). However, the results show that it was not a cultural explanation, but the result of different settings in organizational ecology. Knowing that California could be an exception in terms of its legal system (Gilson, 1998), the results do not change much when excluding California cases. Adding the strength of non-compete clauses per State does not weaken the effect of labor mobility either. As mentioned in Chapter 4, that positive impact of labor mobility is likely to be a result of the concentration of many innovative small firms that stimulates the circulation of skilled engineers. This result is intriguing, and it is suspected that the restriction of non-compete enforcement is less enforceable to small firms, and small firm dominated ecologies. Moreover, we are not sure whether higher mobility rates contribute to more within-industry practices or more across industries practices, something that needs to be explored in future work. In some models, the mediation tests for testing regional mobility rates as the mediator between ecology and commercialization are not very strong. In future work, if we can test the net value of mobile workers by types of knowledge they brought in, we will be better able to clarify the mediation effect argument.

The third research question is to investigate the extent to which the influence of regional ecology on the commercialization propensity varies by firm sizes. Chapter 4 suggests that an ecology dominated with small firms increases the likelihood of commercialization (though not patenting), but the positive effect seems greater for large firms than for small and medium sized firms. As the simulation results show, small and medium sized firms in general generate higher rates of commercialized innovations than large firms do. The results suggest that as the regional ecology becomes more Marshallian-like, the concentration of innovative small firms does not add much value for SMEs. When separating the models by firm size, we find similar results that firm collocations favor large firms over small firms in the region (based on the chi-square test of two models). Furthermore, the results in Chapter 4 also suggest that large firms are not geography-free in terms of searching for local knowledge. However, with limited results, we did not find that large firms capture more skilled inventors in such a region than SMEs since the statistics between the groups is not significant. It cannot be claimed that the power differentials in accessing to skilled inventors favor large more than small firms. However, evidence was found revealing that large firms commercialize more from being in a collaborative R&D project. Similarly, participating in a collaborative R&D project which led to a patent has a negative impact on commercialization for SMEs. If large firms' advantages are to dominate the process of R&D collaborations with local business entities, then it raises policy concerns of power differentials of the ability to involve in the commercialization process between large and small firms. The counter-argument is the division of labor thesis where large and small firms are in charge of different tasks during the innovation process. The current study rejects this argument because in this study the measure of commercialization is a general term that includes in-house commercialization, licensing, and forming start-ups. The negative impact of being in a collaborative project suggests that small firms might not able to receive proper credits. For local policies, the regional innovation system suggests that neighboring members in the system are inevitably connected in different ways, as competitors, suppliers, customers, or collaborators. Failure to recognize the relationship between firm size and differential resources received by members in a region could undercut the regional innovation development. Based on our findings, the policy recommendation is that the government should pay more attention to shape the economic focus with narrow scope. For example, while distributing funding to support technology commercialization and entrepreneurship, the local government should not only focusing on technological fields but also consider the composition of awarded firms against existing regional ecologies.

In line with the goals of this study, the results show that understanding about the ecological characteristics in terms of the regional social and market structures should illustrate a more complete story about the space and innovation relationship. This study also contributes by using nationwide data that increases its generalizability. The findings suggest that the concentration of firms does not always guaranteed co-operations, but depends on how interconnected the firms are in the same region.

5.3 Discussion and research contributions

5.3.1 Specialized vs. diversified

The first part of my dissertation is to clarify the relationship among traditional regional characteristics, regional ecology and regional innovation systems. The findings in this study revisit the agglomeration and regional cluster theory by emphasizing the social structure of firm concentration. The agglomeration economy carries positive external benefits that increase the interaction of communication, the exchange of knowledge sharing, and the trading of technologies between firms (Gertler, 2003).

The findings support Marshall's tradition that the concentration of specialized firms in the same technological field increases regional innovation performance. This finding accords with the argument that the clustering of firms in the same industry increasing competition and "the innovative dynamism arising from it" (Gertler & Wolfe, 2002). The effect of the "local buzz" works better among firms in the same technological domain. Porter's (1990) competitive advantage framework is in this tradition too. In other words, we do not support Jacob's (1969, p129) diversification argument (Feldman and Audretsch, 1999), or the synthetic knowledge argument presented by Asheim (Asheim, 1996).

Why is specialization more important than diversification in patent commercialization? The first conjecture is that the knowledge barriers across different technological fields can be high even for firms located in the same region. Firms in different fields have different professional codes and communication styles, making it difficult to comprehend each other. Secondly, commercialization, in the late stage of innovation, usually requires the participation of related suppliers and service providers (Feldman, 2001). Concerning knowledge searching for commercializing a patent technology, questions need to be more specific and problem solving oriented. Knowledge domains close to each other are more important than distant domains for applicability. The findings suggest that specialization plays a more important role than diversification. If this finding is correct, It is important to consider whether the spillover of mobile skilled workers moved across fields is likely to have more modest effects than of mobile skilled workers moved within fields. We propose an investigation in future research to test the role of regional mobility across and within fields for the innovation process.

5.3.2 Collocation of firms and regional innovation performance: regional ecology explains the knowledge complementarity among firms

This study argues that the concentration of firms of different sizes creates distinctive social structures. We found mixed results for the two innovation performance measures. Firstly, the large firm dominated ecology has a positive impact on regional patenting activities, a finding contradicts the study by Acs et al. (2002), which argues that the presence of large firms decreases regional innovation performance. It can be argued that large firms have changed their patent use strategies in the past twenty years, such as the increase of patenting litigation cases from 1970 to 2000 (Hall and Ziedonis, 2007) that is likely to create incentives for large firms to patent more even if they dominate in the region. However, until we solve the mathematical issue of having counts of all the industrial patents in a MSA as a function of the number of large firm (or small firm) patents in a MSA, we have to be careful claiming our results regarding the relationship between regional ecology and the regional patenting activity.

For the second innovation measure, regional commercialization rates, we found that the small firm dominated ecology increases regional commercialization performance. If large firms indeed use patents for strategic reasons, we should see the concentration of large firms being unable to create a regional innovation system for commercialization because they deny newcomers. By contrast, the increase in small firms not only enhances the local market for ancillary services, but also could reduce entry cost (Vernon, 1969; Chinitz, 1961), which supports the Marshallian thesis.

In addition, this study claims that the regional ecology should be interpreted as a social structure of the space that determines the capacity of multiple knowledge streams; in this case, we focus on the pool of mobile inventors and the R&D capacity of basic research institutions (i.e., local research universities). Our findings support the proposed hypotheses (H2b) that the influence of regional ecology is explained by labor mobility, but not by university knowledge. We found that a SME dominated ecology increases labor mobility and that higher rates of inventor mobility correlate positively with a firm's commercialization propensity. Saxenian's (1996) comparison between Silicon Valley and Route 128 represents two unique regional advantage patterns, which are constructed within different cultural and institutional structures. Our findings suggest that regional ecology is more than a measure of the share of small firm innovations in the locality; it is a measure of regional structure that increases the pool of skilled laborers, and determines interactions between firms in the same region.

5.3.3 Regional ecology and firm size

The second part of this research aims to understand how regional ecologies affect large and small firms differently. Are small firms able to acquire the R&D resources from the spatially integrated institutions to engage in innovation business? Literature on geographic economy mentions that small firms are more likely to rely on external resources in the innovation process than large firms (Feldman, 1994).

This study proposes that the role of large companies is like an innovation hunter instead of an innovation giant or innovation broker. In this study, the findings support the Schumpeterian advantage of large firms being constrained by regional boundaries (Schumpeter, 1942; Galbraith, 1952; Cohen and Klepper, 1996). The first evidence is that large firms benefit positively in an environment with many innovative small firms. The second piece of evidence is that large firms have more in-house commercialization in ecology with many innovative small firms than their neighboring small firms. Thirdly, large firms benefit more from having collaborative partners. In addition, this study presents an unusual finding that the concentration of many innovative small firms has less added value to SMEs. In other words, Hypothesis 3b is rejected. My results show that small firms do not commercialize more patented inventions than large firms in a SME-dominated ecology. Also, the difference in predicted commercialization probability between small firms and large firms is larger in a SME-dominated ecology than in a large-firm dominated ecology. This result is presented in figure 4.2. Contrary to the beliefs of agglomeration economists, a critical nuance that emerges from our findings is that small firms are indifferent to either a Marshallian region or an Anchor-tenant region. This study shows that SMEs are not weaker competitors compared to large firms regardless of the ecological context. For small firms, the simulated predicted probability of commercialization shows no significant variation, neither in a low SME concentration region, nor in a high SME concentration region. In contrast, large firms benefit more from being in a SME dominated region. This study points out a new proposition that large firms are dependent on external knowledge resources as much as small firms. The results also provide some conjecture in terms of large and small firm relationships in a region.

Although we find an ecology dominated by small firms is positive to regional commercialization rates, the data is unable to prove whether this result holds in a pure Marshallian district (regions with an extremely high percentage of small firms). In a sense, it is always a mix of large and small firms in a region and the measure of regional ecology is a ratio measure that describes the relative proportion of organizations in different sizes, ranging from 0% to 67%. Although we do not have pure-Marshallian districts in my data, we do enough variance to show the effect of an additional increase in the percentage of small firm patents on firms' commercialization performance. The findings suggest that when the percentage of small firm innovation increases, large firms are more likely to capture, engage, and appropriate external resources and knowledge than small firms.

In interpreting the results, mechanisms of SME dominated ecologies in enhancing large firm's innovation performance is aligned with the argument of interdependency theory. It suggests that the more small firms are specialized, the more they are engaged in innovation activities. In other words, this also means a larger pool of technology and knowledge being circulated in the region. Hence, SME dominated ecology is a plus for large firms because they view the presence of many innovative SMEs as the great opportunity for acquiring novel technologies and services.

Large firms are likely to be innovation hunters. Although SMEs produce specialized technologies and inventions, they cannot bear the risk of commercializing the invention internally. For survival, SMEs have to seek potential buyers, mostly local buyers. In contrast, large companies are in a privileged position in terms of accessing knowledge and new technologies locally and externally, particularly those multinational corporations. This explains why many large corporations, such as Microsoft (in Mountain View) and IBM (in San Jose), set up branches in Silicon Valley. As the descriptive statistics show, in the triadic patent sample, inventors of large firms are more likely to have multiple patents during a short period (between 2000 and 2003) than small firms. In other words, many small firms usually own one or a few key patents, which directly indicate to their key technologies for market. The Resource-Partitioning theory suggests that when the generalist's market is concentrated, there are niches for specialized entries (Carroll, 1985; Hannan, Polos, and Carroll, 2007). My findings show that when the concentration of large firms in an environment is strong, it provides niches for the neighboring small firms to commercialize their inventions. In other words, small firms outperform large firms in a large-firm dominated ecology. By contrast, when the concentration of small firms is strong, the ecology is crowded with specialized firms and competition between them could be high. Conjecture concerning this finding is that the concentration of many specialized small firms provides large firms opportunities to minimize costs when searching for the optimal complimentary technologies in a small firm dominated ecology. Based on the simulation results, large firm is likely to outperform small firms in a SME-dominated ecology.

Following this thread of discussion, the advantageous role of large firm in a SME dominated region is particularly relevant to policy. An argument similar to the power dynamics thesis claims that large and small firms are having unequal bargaining power, not only technologically, but also politically. For local small firms, the lack of resources cannot only lead to disadvantages in capturing resources from neighbor firms but also can lead to disadvantages when negotiating for policy incentives and favorable regulations, such as innovation agenda, labor skill demands, and funding criteria (Christopherson and Clark, 2007). Christopherson and Clark argue that divergent interests between large corporations and small firms are likely to undercut regional goal. Their concern verifies certain findings in this study that large firms may substantially benefit more from capturing local spillover sources, such as local university knowledge and collaboration opportunities, than small firms. Concerning local policies, the competitive strategy is to recognize the differentials and improve the weakness of innovative small firms, such as their learning and manufacturing capacity, and an increase of incentives to collaborate with local universities.

This study suggests that we have to understand the role of large firms in an innovation ecological system. Large firms are not only able to utilize their internal R&D capacities but are also adept at harnessing the external knowledge sources when the environment provides. In the absence of these privileges, small and medium sized firms have to seek external knowledge sources with more difficulty.

Following the argument from learning region literature, organizational learning and innovation is a geographical process (Scott, 1988; Storper, 1997). This study challenges the existing theory about the presumption of the homogenous social learning process of actors within a region. Particularly, knowledge seeking is a social process. Firms' probability to search external resources are constrained by where they are located and what they are technological capable of (Alcacer and Chung, 2011). In addition to the geographical proximity argument, we found the importance of understanding the organizational heterogeneity and firm composition at the regional level. It can be suggested that the ecological perspective is complementary to understanding the mechanisms of information flow in innovative regions. Future studies should pay more attention to studying these social processes.

5.4 Research limitations

This study has several limitations. First, the use of patent as the proxy for innovation may limit the generalizability of the findings to certain type of industries and firms. In many industries, firms choose different means to protect their intellectual property rights than patent enforcement, such as secrecy, lead-time, complementary manufacturing/services, and other legal approaches (Cohen et al., 2000). Therefore, the interpretation of the findings more accurately represents patent-based industries (e.g., the pharmaceutical industry and computer industry), although they do apply to the patenting strategies of non-patenting industries (e.g., the traditional machinery industries) as well.

Secondly, although the survey instrument is a good tool to collect detailed information about the innovation process in an R&D project, some questions are likely to be too difficult for some respondents (i.e., patent inventors) to answer, such as questions about the commercial use of patents. In some cases, engineers may be involved in the patenting stage and not in the commercialization stage due to the division of labor in the company. Particularly, employees in large firms could be less knowledgeable about the use of a patented invention regarding its downstream commercialization process[11]. However, it is also a tradeoff because inventors are the best candidate to answer other survey question, such as those relating to detailed information about R&D activity, and inventors' career histories and backgrounds.

Thirdly, this study was unable to provide evidence for the network mechanisms embedded in a SME-dominated ecology. Prior research often points out that the concentration of many local small firms is likely to establish a dense network among co-located firms, thus maintaining sustainable long-term inter-firm relationships (Owen-Smith and Powell, 2004). While the findings suggest that the knowledge externality in the presence of many innovative small firms is positive, we did not test whether it was due to the network mechanism at a structural level. However, we would like to explore this mechanism in future work.

Finally, using a survey instrument to investigate whether regional or firm level factors are more important to innovation performance could have a methodological limitation on exploring the research question. However, this study has tried to be distinct from prior research by employing a quantitative methodology using a nationwide dataset with large population to contribute to the regional innovation system.

[11] Respondents of large firms are more likely to report that they did not know the answers to questions related to the use of the patent than those of small firms.

5.5 Future research

As mentioned in the previous section, one limitation of this research uses patents to study the regional innovation system across industries in the United States. This limitation offers an opportunity to expand the research into different country contexts and research avenues. First, we are planning to conduct comparative studies between the United States models and other countries with a similar innovation capacity, for example in Japan and some European countries. Secondly, I would like to construct a formal model for building the theoretical framework to explain the commercialization propensity between large and small firms in the presence of different regional ecologies. Thirdly, I would like to further study the ecological nature of regions and its impact on innovation growth by different innovation measures, as certain researchers argue that patents could be a poor measure of innovation output. Furthermore, for those industries protecting their intellectual properties through other means, such as secrecy, future study can test the role of regional ecology on other type of innovation outputs targeting non-patenting based industries. In addition, another limitation of this study is that the sample was not selecting based on regional distribution. Therefore, we could mainly analyze those regions with at least some innovation. For the future study, I would like to collect new data with better regional coverage to better illustrate ecological distribution across regions.

Furthermore, if a small firm dominated ecology facilitates knowledge flows, what are the processes? I found labor mobility could be one of the mechanisms mediates the positive externality of a SMEs dominated ecology. This study suggests that regional ecology should increase the labor mobility, network density, organizational interactions, and further increase regional innovation performance. We would like to further investigate whether the effect of worker mobility is within or across industry. In other words, it would further verify the specialization versus diversification issue at a regional level. Based on the mobile inventor database that we collected using Lai et al.'s US inventor data, we plan to collect additional information regarding the shift in technological fields after moves. By doing so, we will be better able to clarify the labor mobility mechanism.

As to the social network aspects, Owen-Smith and Powell (2004) suggest that information flows among firms is a function of geographical interpersonal ties and institutional characteristics. Due to data limitations, this study did not have regional network data. I would like to study the role of regional ecology and its impact on the regional networks if network data about collocated firms is available.

Finally, I would like to develop a more nuanced concept of regional ecology in the future. Size concentration of a region is one aspect of the regional structure and a first step of the measure due to data limitations. The findings in this study also point out the discrete needs and beneficiaries among large and small firms, implying that the optimistic prediction of unified shared resources and knowledge spillovers due to regional proximity may not hold. The discussion of competition of co-located firms is under-developed in regional learning literature because it presumes positive externalities while not explicitly explaining the processes and mechanisms. Hence, the regional ecology concept should address competition among firms in the region as well. In short, future work would like to develop relevant theories and measures that might guide future case studies and large-scale studies.

5.6 Policy implications

Inventions and commercialization are the byproduct of a complex innovation process, which is also a geographic process. To conclude, my findings suggest a discrete phenomenon between regional innovation and firm innovation. It has been firstly observed that large firm dominated ecologies enhance regional patenting activities. However, the small firm dominated ecologies increase regional commercialization. At the project level, small firm dominated ecologies have been shown to be beneficial to firm's likelihood to commercialize their patented R&D discoveries. At the regional level, we propose the following policy implications. In particular, state government at the forefront of addressing these issues should pay more attention to the social structure of regions. Firstly, to encourage regional learning, policy makers should not only look at the traditional concentration measures, such as the Herfindahl index or the location Gini index, but also at the ecological structure of the region. The ecological perspective can provide better insights in understanding the discrete needs and benefits of actors like large corporations, small firms, universities, and other institutions, in the same regions. For example, favorable R&D credit from local government is likely to reinforce R&D capacity by attracting more large corporations to the region. As our findings suggest, the presence of anchor tenants in a region increases the patenting activity levels.

Furthermore, concerning regional commercialization performance, policy makers should continue encouraging the influx of small firms in a region, as well as the interconnection of small firms in that region. When it comes to the granting process, funding agencies should allocate funds to local small and medium sized firms in the effort of increasing the specification of technologies. In other words, the support of local business in the conduct of innovation is to increase commercialization and provide incentives to attract skilled and competitive labor. From a policy perspective, keeping small and medium sized enterprises staying in the region regardless of the organizational composition should be the focus of the regional government. Under current debates, small firms are contributing to the creation of local jobs and innovations (Audretsch and Acs, 1991).

We suggest that policy makers and regional researchers should initially understand the structure of the organization ecology in a region. By doing so, they will realize the diversity of environments. Secondly, they should also emphasize heterogeneous capability when conducting R&D and innovation across different types of firms in the same region. In conclusion**,** our policy recommendation reemphasizes the concern of the title of this research, that one size does not fit all. Future policy programs should pay more attention to firm heterogeneity and assuring the competitive advantages of small firms.

Thirdly, by breaking up firm size, I observe that large firms benefit more in the presence of small firm concentration than SMEs do. This study finds that small firms are more innovative than large firms, suggesting that small firms are not suboptimal organizations in the industry in terms of high mortality rates and total production benefits. However, the problem of current policies is they treat regional development as a homogeneous process, and many federal funding programs being underpinned by the assumption that development patterns across regions are unified. For example, the R&D tax credit is one government program that encourages R&D activities. The government also encourages SME innovation through SBIR and STTR grants. We also see various tax credits for venture capital and new business at the state level. Prior studies point out that the impact of R&D tax credit programs led to a lower tax rate for large firms than for small firms (Guenther, 2005), suggesting the benefits do not fill the gaps between large and small firms, something that is likely to discourage the location choice of small firms.

Additionally, small firms in an SME dominated region might have to compete with other local small business in marketing their innovative outputs. This is consistent with the claims of resource partitioning theory that when the environment becomes too crowded with specialists, the entry of another generalist increases its survival rate. By incorporating the regional ecology concept, policy evaluations can investigate regional innovation systems by taking into account the composition of firms in different sizes and types. Additionally, the focus of state policy could investigate if SMEs are disadvantaged when integrating the localized knowledge resources into their R&D system before policy makers implement a variety of policy instruments in regions. For example, local government could ensure the supply of skilled and well-trained workers to the region. For small firms, they may be reluctant to invest in hiring labors with long-term R&D goals. Hence, the local government could assist by providing training or workshops in collaboration with small firms to meet the immediate needs of small firms, such as the programs of Georgia QuickStart organization. Furthermore, local government could also assist with low-cost information about regional economies and marketing data to local business, serving as an information broker.

At the firm level, this study sheds lights on the terms of strategic management by explaining that the acquisition of external knowledge is a geographical process. Each different concentration of firms presents a different type of regional ecology where collocated firms are constrained by the amount of resources they can access. Understanding about the structure of the locality will assist firms to assess the general picture of the complimentary resources they can access, as well as the positive and negative externalities of being in a certain region. For large firms, they should strategically locate themselves in a SME dominated ecology. For small firms, their innovative performance is not significantly affected by the regional ecology they are in, suggesting that they can be innovative in a variety of ecologies. However, our findings also imply that small firms do not significantly benefit from the presence of many innovative small firms. In order to utilize the spillover effects in a SME dominated ecology, it is recommended small firms develop specialized niches to distinguish themselves from other local SMEs.

In this study, inventor mobility and university R&D are considered two important regional resources. Regardless of size, both large and small firms benefit from labor mobility. We did not find significant differences where large firms are benefiting more from the high mobility rates in a region than small firms are. However, university R&D has shown to be more useful (less negative) for large firms than small firms in the same metropolitan areas. As this finding is relatively weak statistically (although the results are consistent along many models), future study should continue examine if there is a power imbalance in the relationship between large and small firms concerning differential access to specialized skill workers and local university resources. In summary, this study provides insights into the mechanisms of information flows caused by the concentration of many small firms in the region and provides firms with useful information to shape their commercialization strategies and innovation management practices.

References

Acs, Z. J., Anselin, L., & Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research Policy*, *31*, 1069-1085.

Afuah, A. (2003). Innovation Management: Strategies, Implementation and Profits. Oxford University Press.

Agrawal, A., & Cockburn, I. (2003). The anchor tenant hypothesis: exploring the role of large, local, R&D-intensive firms in regional innovation systems. *International Journal of Industrial Organization*, *21*(9), 1227-1253.

Agrawal, A., Cockburn, I., & McHale, J. (2006). Gone but not forgotten: knowledge flows, labor mobility, and enduring social relationships. *Journal of Economic Geography*, *6*(5), 571 -591.

Agrawal, A., Cockburn, I., & Rosell, C. (2010). Not Invented Here? Innovation in company towns. *Journal of Urban Economics*, *67*(1), 78-89.

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics Letters*, *80*(1), 123-129.

Alcacer, J., & Chung, W. (2011). Benefiting from location: knowledge retrieval. *Global Strategy Journal*, *1*(3□4), 233-236.

Almeida, P., & Kogut, B. (1999). Localization of Knowledge and the Mobility of Engineers in Regional Networks. *Manage. Sci.*, *45*(7), 905–917. doi:10.1287/mnsc.45.7.905

Asheim, B. T. (1996). Industrial Districts as "Learning Regions": a condition for Prosperity. *European Planning Studies*, *4*(4), 379-400.

Audretsch, D. B., & Acs, Z. J. (1991). Innovation and Size at the Firm Level. *Southern Economic Journal*, *57*(3), 739-744.

Audretsch, D. B., & Feldman, M. P. (1996). R&D Spillovers and the Geography of Innovation and Production. *The American Economic Review*, *86*(3), 630-633.

Audretsch, D. B., & Stephan, P. E. (1996). Company-Scientist Locational Links: The Case of Biotechnology. *American Economic Review*, *86*(3), 641-52.

Boschma, R. (2005). Editorial: Role of Proximity in Interaction and Performance: Conceptual and Empirical Challenges. *Regional Studies*, *39*(1), 41-45.

Boschma, R., & Frenken, K. (2009). *Technological relatedness and regional branching*. Utrecht University, Section of Economic Geography.

Breschi, S., & Lissoni, F. (2009). Mobility of skilled workers and co-invention networks: an anatomy of localized knowledge flows. *Journal of Economic Geography*, *9*(4), 439-468.

Bresnahan, T., & Gambardella, A. (Eds.). (2004). *Building High-Tech Clusters: Silicon Valley and Beyond*. Cambridge University Press.

Breznitz, D., & Taylor, M. (2009). The Communal Roots of Entrepreneurial-Technological Growth? Social Fragmentation and the Economic Stagnation of Atlanta's Technology Cluster. *SSRN eLibrary*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1481857

Breznitz, S. M., & Anderson, W. P. (2005). Boston metropolitan area biotechnology cluster. *Canadian Journal of Regional Science*, *XXVIII*(2), 249-264.

Carroll, G. R. (1985). Concentration and Specialization: Dynamics of Niche Width in Populations of Organizations. *American Journal of Sociology*, *90*(6), 1262-1283.

Carroll, G. R., & Swaminathan, A. (2000). Why the Microbrewery Movement? Organizational Dynamics of Resource Partitioning in the U.S. Brewing Industry. *American Journal of Sociology*, *106*(3), 715-762.

Casper, S. (2007). How do technology cluster emerge and become sustainable? Social network formation and inter-firm mobility within the San Diego biotechnology cluster. *Research Policy*, *36*, 438-455.

Chesbrough, H. W. (2003). Open innovation: the new imperative for creating and profiting from technology. Harvard Business Press.

Christopherson, S., & Clark, J. (2007). Power in firm network: What it means for regional innovation systems. *Regional Studies*, *41*, 1223-1236.

Clark, J., Huang, H.-I., & Walsh, J. P. (2010). A typology of "innovation districts": what it means for regional resilience. *Cambridge J Regions Econ Soc*, *3*(1), 121-137.

Cohen, W. M., & Klepper, S. (1996). Firm Size and the Nature of Innovation within Industries: The Case of Process and Product R&D. *The Review of Economics and Statistics*, *78*(2), 232-243.

Cohen, W. M., & Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, *35*(1), 128-152.

Cohen, W. M., Levin, R. C., & Mowery, D. C. (1987). Firm Size and R&D Intensity: A Re-Examination. *National Bureau of Economic Research Working Paper Series*, *No. 2205*. Retrieved from http://www.nber.org/papers/w2205

Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2000). Protecting Their Intellectual Assets: Appropriability Conditions and Why U.S. Manufacturing Firms Patent (or Not). *National Bureau of Economic Research Working Paper Series*, *No. 7552*. Retrieved from http://www.nber.org/papers/w7552

Cohen, W. M., Nelson, R. R., & Walsh, J. P. (2002). Links and Impacts: The influence of public research on industrial R&D. *Management science*, *48*(1), 1-23.

Cooke, P., & Morgan, K. (1994). The regional innovation system in Baden-Württemberg. *Journal of Technology Management*, *9*, 394-429.

Crevoisier, O. (2009). The Innovative Milieus Approach: Toward a Territorialized Understanding of the Economy? *Economic Geography*, *80*(4), 367-379.

Delgado, M., Porter, M. E., & Stern, S. (2010). Clusters and entrepreneurship. *Journal of Economic Geography*, *10*(4), 495 -518.

Fallick, B., Fleischman, C. A., & Rebitzer, J. B. (2011). Job-Hopping in Silicon Valley: Some Evidence Concerning the Microfoundations of a High-Technology Cluster. *Review of Economics and Statistics*, *88*(3), 472-481.

Feldman, M. P. (1994). Knowledge complementarity and innovation. *Small Business Economics*, *6*(5), 363-372.

Feldman, M. P. (2001). The Entrepreneurial Event Revisited: Firm Formation in a Regional Context. *Industrial and Corporate Change*, *10*(4), 861-891. doi:10.1093/icc/10.4.861

Feldman, M. P., & Audretsch, D. B. (1999). Innovation in cities:: Science-based diversity, specialization and localized competition. *European Economic Review*, *43*(2), 409-429. doi:10.1016/S0014-2921(98)00047-6

Feldman, M. P., & Kogler, D. F. (2010). Chapter 8 Stylized Facts in the Geography of Innovation. *Handbook of The Economics of Innovation* (Vol. 1). Oxford: Elservier B.V.

Fernandez-Ribas, A., & Shapira, P. (2009). The role of national and regional innovation programmes in stimulating international cooperation in innovation. *International Journal of Technology Management*, *48*(4), 473 - 498.

Fleming, L. (2001). Recombinant Uncertainty in Technological Search. *Management Science*, *47*(1), 117-132.

Florida, R. (1995). Toward the learning region. *Futures*, *27*(5), 527-536.

Gambardella, A., Harhoff, D., & Verspagen, B. (2008). The Value of European Patents. *European Management Review*, *5*(2), 69-84.

Garmaise, M. J. (2009). Ties that Truly Bind: Noncompetition Agreements, Executive Compensation, and Firm Investment. *Journal of Law, Economics, and Organization*.

Gertler, M. S. (2003). Tacit knowledge and the economic geography of context, or The undefinable tacitness of being (there). *Journal of Economic Geography*, *3*, 75-99.

Gertler, M. S., & Wolfe, D. A. (2002). Innovation and social learning: institutional adaptation in an era of technological change. New York: Palgrave Macmillan.

Gilson, R. J. (1998). The Legal Infrastructure of High Technology Industrial Districts: Silicon Valley, Route 128, and Covenants Not to Compete. *SSRN eLibrary*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=124508

Glaeser, E. L., Kallal, H. D., Scheinkman, J. A., & Shleifer, A. (1992). Growth in Cities. *Journal of Political Economy*, *100*(Centennial issue), 1126-1152.

Glaeser, E. L., Scheinkman, J., & Shleifer, A. (1995). Economic growth in a cross-section of cities. *Journal of Monetary Economics*, *36*(1), 117-143.

Gordon, L. R., & McCann, P. (2000). Industrial Clusters: Complexes, Agglomeration and/or Social Networks? *Urban Studies*, *37*(3), 513-532.

Griliches, Z. (1979). Issues in Assessing the Contribution of Research and Development to Productivity Growth. *Bell Journal of Economics*, *10*(1), 92-116.

Griliches, Z. (1991). *Patent Statistics as Economic Indicators: A Survey* (NBER Working Paper No. 3301). National Bureau of Economic Research, Inc. Retrieved from http://ideas.repec.org/p/nbr/nberwo/3301.html

Guo, G., & Zhao, H. (2000). Multilevel Modeling for Binary Data. *Annual Review of Sociology*, *26*, 441-462.

Hagedoorn, J. (2003). Sharing intellectual property rights—an exploratory study of joint patenting amongst companies. *Industrial and Corporate Change*, *12*(5), 1035 -1050.

Hall, B., Jaffe, A., & Trajtenberg, M. (2001). The NBER patent citations data file: Lessons, insights, and methodological tools. Retrieved from

Haltiwanger, J. C., Jarmin, R. S., & Miranda, J. (2010). *Who Creates Jobs? Small vs. Large vs. Young* (NBER Working Paper No. 16300). National Bureau of Economic Research, Inc. Retrieved from http://ideas.repec.org/p/nbr/nberwo/16300.html

Hannan, M. T., & Freeman, J. (1977). The Population Ecology of Organizations. *American Journal of Sociology*, *82*(5), 929-964.

Hannan, M. T., Polos, L., & Carroll, G. R. (2007). *Logics of Organization Theory: Audiences, Codes, and Ecologies*. Princeton University Press.

Harrison, B. (1992). Industrial Districts: Old Wine in New Bottles? *Regional Studies*, *26*(5), 469-483.

Hicks, D., Breitzman, T., Olivastro, D., & Hamilton, K. (2001). The changing composition of innovative activity in the US — a portrait based on patent analysis. *Research Policy*, *30*(4), 681-703.

Hoetker, G. (2007). The use of logit and probit models in strategic management research: Critical issues. *Strategic Management Journal*, *28*(4), 331-343.

Hoisl, K. (2007). Tracing mobile inventors - The causality between inventor mobility and inventor productivity. *Research Policy*, *36*, 619-636.

Hoisl, K. (2009). Does mobility increase the productivity of inventors? *The Journal of Technology Transfer*, The Journal of Technology Transfer, *34*(2), 212-225.

Jacobs, J. (1969). *The economy of cities*. Random House.

Jaffe, A. (1998). The importance of "Spillover" in the policy mission of the Advanced Technology Program. *Journal of Technology Transfer*, *23*(2), 11-19.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*, *108*(3), 577-598.

Kenney, M., & Florida, R. (1994). The organization and geography of Japanese R&D: results from a survey of Japanese electronics and biotechnology firms. *Research Policy*, *23*(3), 305-322.

King, G., Tomz, M., & Wittenberg, J. (2000). Making the most of statistical analyses: Improving interpretation and presentation. *AMERICAN JOURNAL OF POLITICAL SCIENCE*, 341--355.

Kremp, E., & Mairesse, J. (2004). *Knowledge Management, Innovation, and Productivity: A Firm Level Exploration Based on French Manufacturing CIS3 Data* (NBER Working Paper No. 10237). National Bureau of Economic Research, Inc. Retrieved from http://ideas.repec.org.www.library.gatech.edu:2048/p/nbr/nberwo/10237.html

Lai, R., D'Amour, A., & Fleming, L. (2009). *The careers and co-authorship networks of U.S. patent-holders, since 1975*. Harvard Institute for Quantitative

Social Science. Retrieved from http://dvn.iq.harvard.edu/dvn/dv/patent/faces/study/StudyPage.xhtml?studyId=38083

Lenzi, C. (2009). Workers' mobility and patterns of knowledge diffusion: evidence from Italian data. *The Journal of Technology Transfer*, *35*, 651-670.

Levitt, B., & March, J. G. (1988). Organizational Learning. *Annual Review of Sociology*, *14*(1), 319-338.

Long, S. J. (2009). Group comparisons in logit and probitusing predicted probabilities. Indiana University.

MacKinnon, D. P., & And Others. (1995). A Simulation Study of Mediated Effect Measures. *Multivariate Behavioral Research*, *30*(1), 41-62.

Macdonald, S. (1986). The distinctive research of the individual inventor. *Research Policy*, *15*(4), 199-210.

Markusen, A. (1996). Sticky places in slippery space: A typology of industrial districts. *Economic Geography*, *72*(3), 293-313.

Marshall, A. (1920). *Principles of Economics*. (8th, Ed.). Macmilan, London.

Marx, M., Strumsky, D., & Fleming, L. (2009). Mobility, Skills, and the Michigan Non-Compete Experiment. *MANAGEMENT SCIENCE*, *55*(6), 875-889.

Mattes, E., Stacey, M. C., & Marinova, D. (2006). Surveying inventors listed on patents to investigate determinants of innovation. *Scientometrics*, *69*, 475-498.

McCullagh, P., & Nelder, J. A. (1989). *Generalized Linear Models*. CRC Press.

McPherson, M. (1983). An Ecology of Affiliation. *American Sociological Review*, *48*(4), 519-532.

Motohashi, K. (2005). University–industry collaborations in Japan: The role of new technology-based firms in transforming the National Innovation System. *Research Policy*, *34*(5), 583-594.

Motohashi, K., & Muramatsu, S. (2012). Examining the university industry collaboration policy in Japan: Patent analysis. *Technology in Society*, *34*(2), 149-162.

Mowery, D. C. (1998). The changing structure of the US national innovation system: implications for international conflict and cooperation in R&D policy. *Research Policy*, *27*(6), 639-654.

Nelson, R. (2001). Evolutionary Perspective on Economic Growth, *74*(3), 165-194.

Nelson, R. R., & Winter, S. G. (1982). *An Evolutionary Theory of Economic Change*. Belknap Press of Harvard University Press.

Nonaka, I., & Takeuchi, Hi. (1995). *The knowledge-creating company*. New York: Oxford University Press.

Oden, M. (2000). Federal Defense Industrial Policy, Firm Strategy, and Regional Conversion Initiatives in Four American Aerospace Regions. *International Regional Science Review*, *23*(1), 25-47.

Oettl, A., & Agrawal, A. (2008). International labor mobility and knowledge flow externalities. *Journal of International Business Studies*, *39*(8), 1242-1260.

Owen-Smith, J., & Powell, W. W. (2004). Knowledge Networks as Channels and Conduits: The Effects of Spillovers in the Boston Biotechnology Community. *Organization Science* (pp. 5-21).

Pavitt, K. (1991). Key Characteristics of the Large Innovating Firm*. *British Journal of Management*, *2*(1), 41-50.

Piore, M., & Sabel, C. (1986). The Second Industrial Divide: Possibilities For Prosperity. Basic Books.

Porter, M. (1990). The competitive advantage of Nations.

Powell, W. W. (1990). Neither Market Nor Hierarchy: Network Forms of Organization. *Research In Organizational Behavior*, *12*, 295-336.

Pyke, F., & Sengenberger, W. (Eds.). (1992). *Industrial districts and local economic regeneration*. Geneva: International Institute for Labour Studies.

Rahm, D., Kirkland, J., & Bozeman, B. (1999). *University-Industry R&D Collaboration in the United States, the* (1st ed.). Springer.

Raudenbush, S. W., & Bryk, A. S. (2002). Hierarchical Linear Models: Applications and Data Analysis Methods. SAGE.

Rosenkopf, L., & Almeida, P. (2003). Overcoming local search through alliances and mobility. *Management science*, *49*(6), 751-766.

Rothaermel, F. T., & Deeds, D. L. (2004). Exploration and exploitation alliances in biotechnology: A system of new product development. *Strategic Management Journal*, *25*, 201-221.

Saxenian, A. (1996). *Regional Advantage*. Cambridge: Harvard University Press.

Schimank, U. (1988). The contribution of university research to the technological innovation of the German economy: Societal auto-dynamic and political guidance. *Research Policy*, *17*(6), 329-340.

Schumpeter, J. A. (1942). *Capitalism, Socialism and Democracy*. New York, London: Harper & Brothers.

Singh, J., & Agrawal, A. (2011). Recruiting for Ideas: How Firms Exploit the Prior Inventions of New Hires. *Management Science*, *57*(1), 129-150.

Singh, J., & Marx, M. (2011). The Geographic Scope of Knowledge Spillovers: Spatial Proximity, Political Borders and Non-Compete Enforcement Policy. *SSRN eLibrary*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1541794

Sobel, M. E. (1982). Asymptotic Confidence Intervals for Indirect Effects in Structural Equation Models. *Sociological Methodology*, *13*, 290-312.

Song, J., Almeida, P., & Wu, G. (2003). Learning-by-hiring: When is mobility more likely to facilitate interfirm knowledge transfer? *Management science*, *49*(4), 351-365.

Stokes, D. (1997). *Pasteur's Quadrant*. Washington DC: Brookings.

Stolpe, M. (2001). Mobility of Research Workers and Knowledge Diffusion as Evidenced in Patent Data - The Case of Liquid Crystal Display Technoloy. Retrieved from

Storper, M. (1997).The regional world: Territorial development in a global economy. New York, NY: Guildford Press.

Storper, M., & Venables, A. J. (2004). Buzz: face-to-face contact and the urban economy. *Journal of Economic Geography*, *4*(4), 351 -370. doi:10.1093/jnlecg/lbh027

Stuart, T., & Sorenson, O. (2003). The geography of opportunity: spatial heterogeneity in founding rates and the performance of biotechnology firms. *Research Policy*, *32*(2), 229-253.

Teece, D. J. (1986). Profiting from technological innovation: Implications for integration, collaboration, licensing and public policy. *Research Policy*, *15*(6), 285-305.

Thompson, P., & Kean, M. F. (2005). Patent Citations and the Geography of Knowledge Spillovers: A Reassessment. *American Economic Review*, *95*(1), 450-460.

Walsh, J. P. (1993). Supermarkets Transformed: Understanding Organizational and Technological Innovations (illustrated edition.). Rutgers University Press.

Youtie, J., & Shapira, P. (2008). Building an innovation hub: A case study of the transformation of university roles in regional technological and economic development. *Research Policy*, *37*, 1188-1204.

Zaheer, A., & Hernandez, E. (2011). The geographic scope of the MNC and its alliance portfolio: Resolving the paradox of distance. *Global Strategy Journal*, *1*(1□2), 109-126.

Zelner, B. A. (2009). Using simulation to interpret results from logit, probit, and other nonlinear models. *Strategic Management Journal*, *30*(12), 1335-1348.

Zucker, L. G., & Darby, M. R. (1996). Star scientists and institutional transformation: Patterns of invention and innovation in the formation of the biotechnology industry. *PNAS*, *93*, 12709-12716.

Zucker, L. G., Darby, M. R., & Torero, M. (2002). Labor Mobility from Academe to Commerce. *Journal of Labor Economics*, Journal of Labor Economics, *20*(3), 629-660.